

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
28 November 2002 (28.11.2002)

PCT

(10) International Publication Number  
**WO 02/095007 A2**

(51) International Patent Classification<sup>7</sup>: **C12N**

(21) International Application Number: PCT/US02/16819

(22) International Filing Date: 23 May 2002 (23.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/293,267 23 May 2001 (23.05.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:  
US 60/293,267 (CIP)  
Filed on 23 May 2001 (23.05.2001)

(71) Applicant (for all designated States except US): **CORVAS INTERNATIONAL, INC.** [US/US]; 3030 Science Park Road, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MADISON, Edwin, L.** [US/US]; 11005 Cedarcrest Way, San Diego, CA 92121 (US). **SEMPLE, Joseph, Edward** [US/US]; 9711 Caminito Pudregak, San Diego, CA 92131 (US). **VLA-SUK, George, P.** [US/US]; 7325 Calle Luna, Carlsbad, CA 92009 (US). **KEMP, Scott, Jeffrey** [US/US]; 7873 Avenida Navidad, #263, San Diego, CA 92122 (US). **KOMANDLA, Mallareddy** [IN/US]; 8148 Genessee Avenue, #30, San Diego, CA 92122 (US). **SIEV, Daniel, Vanna** [US/US]; 10415 Westchester Avenue, San Diego, CA 92126 (US).

(74) Agents: **SEIDMAN, Stephanie, L.** et al.; Heller Ehrman White & McAuliffe LLP, 4350 La Jolla Village Drive, San Diego, CA 92122-1246 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PI, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/095007 A2

(54) Title: CONJUGATES ACTIVATED BY CELL SURFACE PROTEASES AND THERAPEUTIC USES THEREOF

(57) Abstract: Conjugates, compositions and method for treatment, prevention, or amelioration of one or more symptoms of cell surface protease-related diseases, including MTSP-related, urokinase-type plasminogen activator (uPA) or endotheliase-related diseases, are provided. The conjugates for use in the compositions and methods are peptidic conjugates that contain therapeutic, including cytotoxic, agents.

-1-

**CONJUGATES ACTIVATED BY CELL SURFACE PROTEASES AND  
THERAPEUTIC USES THEREOF**

**RELATED APPLICATIONS**

Benefit of priority to U.S. provisional application Serial No. 60/293,267, filed May 23, 2001, to Edwin L. Madison, Joseph Edward Semple and George P. Vlasuk, entitled "CONJUGATES ACTIVATED BY CELL SURFACE PROTEASES  
5 AND THERAPEUTIC USES THEREOF" is claimed. Where permitted, the subject matter of the application is incorporated by reference in its entirety.

**FIELD OF THE INVENTION**

Conjugates, compositions and methods for localized delivery of therapeutic agents for treating a variety of disorders, such as , proliferative  
10 diseases, autoimmune diseases, infectious diseases and inflammatory diseases, are provided. The conjugates, which act as prodrugs, contain therapeutic agents and peptidic substrates that are cleaved by cell surface proteases to release therapeutic agents in the vicinity of the targeted cells.

**BACKGROUND**

15 Effective treatment of cancer and other proliferative diseases involves administration of chemotherapeutic agents, typically systemic administration. Typically chemotherapeutic agents are cytotoxic agents that act by inhibiting proliferation or other metabolic processes, so that actively proliferating and growing cells will be targeted by the agent. Such targeting, however, is not  
20 highly specific, and the side-effects are often devastating.

Thus, a goal in pharmacology is the design of specific agents that act with high specific activity on targeted cells or tissues. This aim is of particular importance, for example, in the design of agents for treatments of diseases, such as proliferative diseases, including neoplastic disease, and diseases of viral  
25 origin, in which the ratio of toxic dose to therapeutic dose is generally close to one and the dosage must be restricted. Numerous approaches to achieving this goal have been developed. Among these are the use of conjugates that contain a targeting agent, such as an antibody and/or growth factor, and a therapeutic agent, that act on specific cells; the use of antisense technology that is targeted  
30 to specific genes and/or proteins; the use of genetic therapy to provide, for

-2-

example, correct copies of defective genes or pharmaceutically active compounds, and the use of toxins that are relatively non-toxic unless delivered intracellularly. Thus far success has been limited. There are only a limited number and type of potential targeting agents, and the specificity of such agents is optimal.

Hence there is a need to develop means for delivery of therapeutic agents to targeted cells and tissues. Therefore, it is an object herein, among others, to provide methods and compounds for targeted delivery of therapeutic agents.

#### SUMMARY OF THE INVENTION

Provided herein are compounds and methods for targeted delivery of therapeutic agents. The compounds are conjugates that contain a peptidic substrate for a cell surface protease, or a soluble, shed or released form thereof, and an agent that upon cleavage by the protease is a therapeutic agent or in a form that can be activated by the targeted cell or tissue or in the local environment thereof. The agents include therapeutic agents, such as cytotoxic agents, drugs, therapeutic nucleic acid molecules, and diagnostic agents, such as labelled moieties and imaging agents. The cell surface proteases are proteases located at a cell surface and, include, but are not limited to, membrane-bound proteases such as membrane-bound serine proteases (SPs), including, for example, proteases designated MTSPs and endotheliasins. Also contemplated are proteases that are located at the cell surface by virtue of a specific binding interaction with a receptor therefor. Included among such proteases is urokinase plasminogen activator (u-PA; see, *e.g.*, Hung (1984) *Adv. Exp. Med. Biol.* 172:281-293; Cheng *et al.* (1989) *Gene* 69:357-363) bound to urokinase plasminogen activator receptor (u-PAR). The conjugates contain one or more substrates for one or a plurality of cell surface proteases linked either directly or via a linker to a targeted agent, including a therapeutic agent, such as a cytotoxic agent. The conjugates provided herein contain the following components: (peptidic substrate)<sub>s</sub>, (linker)<sub>q</sub>, and (targeted agent)<sub>t</sub>, in which: at least one peptidic substrate moiety is linked with or without a linker (L) to at least one therapeutic agent, s is 1 or more and each substrate is the same or different, and is typically is between 1 and 6, generally 1, 2 or 3; q is 0 or more

-3-

as long as cell surface protease(s) cleaves the peptidic substrate(s) and releases active therapeutic agent or, releases the agent in a form that is converted by the cell, tissue or surrounding environment to an active form, q is 0 to t, generally 1 to 4; t is 1 or more, generally 1 or 2 and each targeted agent are the same or different; linker refers to any linker; and the targeted agent is any agent, typically a therapeutic agent, such as a cytotoxic agent, a nucleic acid, a diagnostic agent, such as an imaging agent or labeled moiety, or a drug, including, but not limited to, anti-tumor, anti-cancer, anti-angiogenic, pro-apoptotic and anti-mitotic agents or treatments.

10 The therapeutic agents include any biologically active molecule. These agents include toxins, cytokines and lymphokines, growth factors, nucleic acid molecules, such as antisense nucleic acid, dsRNA, and DNA molecules. The therapeutic agents include those that are active intracellularly, such as cytotoxins, or extracellularly, such as modulators of the activity of extracellular  
15 receptors. When in the conjugates the therapeutic agents are substantially inactive, and when cleaved are released in active form or in a form that can be activated by the targeted cell or tissue or environment thereof.

In an exemplary embodiment, the conjugates for use in the methods and compositions provided herein can be represented by the formula:

20  $(\text{peptide}^i)_s-(\text{linker})_q-(\text{therapeutic agent})_t$   
or a derivative thereof, where peptide<sup>i</sup> is a peptidic substrate for a cell surface protease; s is greater than or equal to 1, or is 1 to 6, or is 1 or 2, or is 1; linker is any linker; q is greater than or equal to 0, or is 0 to 4, or is 0 or 1; the therapeutic agent is, for example, a cytotoxic agent, including, but not limited  
25 to, an anti-tumor, anti-angiogenic, anti-cancer, pro-apoptotic and anti-mitotic agents; and t is 1 or more, or is 1 or 2. In these conjugates, the therapeutic agent is covalently attached, optionally via a linker L, to either the C-terminus or the N-terminus of the peptidic substrate.

In certain embodiments, peptide<sup>i</sup> is a substrate for a cell surface protease  
30 whereby, upon action of the protease, the conjugate, which is substantially inactive, is cleaved at a point on the peptidic substrate chain to release a compound of the formula:



-4-

(peptide<sup>a</sup>)<sub>s</sub>-(linker)<sub>q</sub>-(therapeutic agent)<sub>t</sub>

or a derivative thereof, that exhibits therapeutic activity *in vitro* and/or *in vivo*.

In these conjugates, the therapeutic agent is, for example, a cytotoxic agent,

and peptide<sup>a</sup> is a truncated version of peptide<sup>1</sup> resulting from cleavage at the P1-

5 P1' bond.

The conjugates can be used to target and deliver the targeted agents to specific cells, and hence can be used for the treatment any diseases that are associated with cells or tissues that express a cell surface protease, including cell-associated and cell-localized proteases. The cells on which or near which  
10 such proteases are expressed are not necessarily involved in the disease or disease process, but are present and can serve to present the protease, which cleaves the targeted conjugate.

Methods of treatment of diseases associated with cells or tissues that express a cell surface protease, including cell-associated and cell-localized  
15 proteases. The diseases include, but not limited to, proliferative diseases, autoimmune diseases, infectious diseases and inflammatory diseases. For example, diseases include e, but are not limited to, rheumatoid arthritis, lupus, multiple sclerosis, psoriasis, diabetic retinopathies, other ocular disorders, including recurrence of pterygii, scarring excimer laser surgery and glaucoma  
20 filtering surgery, various disorders of the anterior eye, cardiovascular disorders, restenosis, chronic inflammatory diseases, wounds, circulatory disorders, crest syndromes, bacterial infections, viral diseases, including AIDS, dermatological disorders, and cancer, including solid neoplasms and vascular tumors, including, but are not limited to, lung, colon, esophageal, breast, ovarian and prostate  
25 cancers.

Also provided are methods for identifying proteases to target conjugates for treatment of diseases. The methods involve identifying cell-surface protease-associated disease by identifying a cell involved in the disease process or a cell in the vicinity of the cell involved in the disease process; and  
30 identifying a cell surface protease on the cell. Conjugates that target such proteases as provided herein can then be prepared.

**DESCRIPTION OF THE FIGURES**

Figures 1-5 provide *in vitro* CT<sub>50</sub> (time for 50% cleavage) (min) for exemplary conjugates provided herein: A = 0.1-25 min; B = 25-100 min; C = 100-250 min; D = > 250 min.

5        Figure 1 shows exemplary doxorubicin conjugates provided herein and *in vitro* CT<sub>50</sub> (min) data for cleavage of the conjugates by MTSP1.

Figure 2 shows exemplary doxorubicin conjugates provided herein and *in vitro* CT<sub>50</sub> (min) data for cleavage of the conjugates by u-PA.

10       Figure 3 shows exemplary taxol conjugates provided herein and *in vitro* CT<sub>50</sub> (min) data for cleavage of the conjugates by MTSP1.

Figure 4 shows exemplary taxol conjugates provided herein and *in vitro* CT<sub>50</sub> (min) data for cleavage of the conjugates by u-PA.

15       Figure 5 shows exemplary doxorubicin and taxol conjugates provided herein and *in vitro* CT<sub>50</sub> (min) data for cleavage of the conjugates by ET1 (endotheliase 1).

**DETAILED DESCRIPTION OF EMBODIMENTS****A. Definitions**

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the invention(s) belong. All patents, patent applications, published applications and publications, Genbank sequences, websites and other published materials referred to throughout the entire disclosure herein, unless noted otherwise, are incorporated by reference in their entirety. In the event that there are a plurality of definitions for terms herein, those in this section prevail.

20       Where reference is made to a URL or other such identifier or address, it is understood that such identifiers can change and particular information on the internet can come and go, but equivalent information can be found by searching the internet. Reference thereto evidences the availability and public dissemination of such information.

30       As used herein, a targeted agent is any agent intended for targeted delivery and includes therapeutic agents and diagnostic agents and any other agent intended for targeted delivery.

-6-

As used herein, targeted delivery means delivery to a selected cell or tissue that expresses a protease that releases the targeted agent. Such delivery does not have to be exclusively to such selected cell or tissue, but must include it, and generally delivers higher amounts to such selected cells or tissues.

- 5 Delivery includes introduction into a cell or tissue or binding to the cell or tissue or release in the vicinity of the cell or tissue. For example, in some instances, a tumor induces production of proteases, receptors, co-factors or substrates associated with the stroma; delivery, thus, includes targeting such induced stromal activities, such as proteases, receptors and/or enzyme co-factors, in
- 10 invading cells or cells in the tumor that is targeted.

As used herein, therapeutic index is the ratio of  $LD_{50}/ED_{50}$ .

- As used herein, a therapeutic agent is any drug or other agent that is intended for delivery to a targeted cell or tissue, such as proliferating cells, including tumor cells and cells involved in a proliferative, typically an
- 15 undesirable, response. Therapeutic agents, include, but are not limited to, anti-cancer agents, anti-angiogenic agents, pro-apoptotic agents, anti-mitotic growth factors, cytokines, such as tumor necrosis factors and interleukins, and cytotoxic agents and other such agents as described herein and known to those of skill in the art. Therapeutic agents include those that are active upon
- 20 internalization and also those that act extracellularly, such modulators of the activities of certain cell surface receptors, such as G proteins that transduce extracellular signals.

- As used herein, an inactive therapeutic agent is a therapeutic agent that is conjugated to a peptide and thereby, either by virtue of conformational
- 25 changes or size or other factors such as steric hinderance does not exhibit any or exhibits substantially reduced activity compared to the released active therapeutic agent. For example, conjugated doxorubicin is not toxic to cells until it is released from the conjugate in a form that can enter the cell. Upon cleavage of the agent from the conjugate it is in active form or in a form that is
- 30 further processed by one or a plurality of steps, including enzymatically or chemically, in or on the cell, into an active form.

-7-

As used herein, an active therapeutic agent is a therapeutic agent that has been released from the conjugate by cleavage of the peptidic substrate portion of the conjugate. The active therapeutic agent is by virtue of cleavage able to exhibit its intended activity, typically by entering the cell. When  
5 conjugated the therapeutic agents have reduced or no activity as therapeutic agents, and upon cleavage are released in the vicinity of a cell.

As used herein, an anti-cancer agent (used interchangeably with "anti-tumor or anti-neoplasm agent") refers to any agents used in the treatment of cancer. These include any agents, when used alone or in combination with  
10 other compounds, that can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission of clinical symptoms or diagnostic markers associated with neoplasm, tumor or cancer, and can be used in methods, combinations and compositions provided herein. Non-limiting examples of anti-neoplasm agents include anti-angiogenic agents, alkylating agents,  
15 antimetabolite, certain natural products, platinum coordination complexes, anthracenediones, substituted ureas, methylhydrazine derivatives, adrenocortical suppressants, certain hormones, antagonists and anti-cancer polysaccharides.

As used herein, substantially inactive with reference to the conjugated therapeutic agent means at least 1%, generally 10, 20, 30, 50, 60, 70, 80 or  
20 90 or 100% inactive compared to the unconjugated therapeutic agent in a standard or art-recognized assays, such as *in vitro* or *in vivo* assays, that assess the therapeutic activity of the agent.

As used herein, a targeted cell or tissue refers to the cells or tissues that include cell surface proteases that cleave the conjugates. The cells or tissues  
25 can be involved in the disease or can be present at the disease loci or locus by virtue of participation in the disease process or merely serendipitously.

As used herein, angiogenesis is intended to broadly encompass the totality of processes directly or indirectly involved in the establishment and maintenance of new vasculature (neovascularization), including, but not limited  
30 to, neovascularization associated with tumors.

As used herein, anti-angiogenic treatment or agent refers to any therapeutic regimen and compound, that, when used alone or in combination

-8-

with other treatment or compounds, can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission, one or more clinical symptoms or diagnostic markers associated with undesired and/or uncontrolled angiogenesis.

Thus, for purposes herein an anti-angiogenic agent refers to an agent that

- 5 inhibits the establishment or maintenance of vasculature. Such agents include, but are not limited to, anti-tumor agents, and agents for treatments of other disorders associated with undesirable angiogenesis, such as diabetic retinopathies, hyperproliferative disorders and others.

- 10 As used herein, non-anti-angiogenic anti-tumor agents refer to anti-tumor agents that do not act primarily by inhibiting angiogenesis. Whether anti-tumor agents act primarily by inhibiting angiogenesis can be determined using the assays provided herein, or using other assays well known to those of skill in the art.

- 15 As used herein, undesired and/or uncontrolled angiogenesis refers to pathological angiogenesis wherein the influence of angiogenesis stimulators outweighs the influence of angiogenesis inhibitors. As used herein, deficient angiogenesis refers to pathological angiogenesis associated with disorders where there is a defect in normal angiogenesis resulting in aberrant angiogenesis or an absence or substantial reduction in angiogenesis.

- 20 As used herein, a cell surface protease is any protease that is located on or at a cell surface and/or proteases that are located at the cell surface by virtue of a specific binding interaction with a receptor therefor, or that is localized at or near or associated with the cell surface. An exemplary protease located at the cell surface by virtue of a specific binding interaction with a receptor therefor is
- 25 urokinase plasminogen activator (u-PA) bound to urokinase plasminogen activator receptor (u-PAR). Hence cell surface proteases contemplated herein include cell surface-associated proteases. It also includes all forms thereof that can be circulating or inside a cell. To be categorized as a cell surface protease, there must be at least one form thereof that is located (*i.e.* on the surfaces such
- 30 as transmembrane protease or bound to receptor therefor) on the surface of a cell at some point in its cycle. Cell surface protease include serine proteases,

such as, but are not limited to, the transmembrane serine protease (MTSPs) and endotheliases and urokinases.

As used herein, a serine protease (SP) refers to a diverse family of proteases in which a serine residue is involved in the hydrolysis of proteins or peptides. The serine residue can be part of the catalytic triad mechanism, which includes a serine, a histidine and an aspartic acid in the catalysis, or be part of the hydroxyl/ $\epsilon$ -amine or hydroxyl/ $\alpha$ -amine catalytic dyad mechanism, which involves a serine and a lysine in the catalysis. Of particular interest are SPs of mammalian, including human, origin. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al.* (1987) *Molecular Biology of the Gene*, 4th Edition, The Benjamin/Cummings Pub. co., p.224).

As used herein shed, soluble and released forms of cell surface proteases are contemplated. Such forms include, for example, forms found in serum upon proteolytic degradation or other removal of the extracellular portion of membrane bound protease, and splice variants that do not include a transmembrane domain.

As shown herein, the protease activity of cell surface proteases and proteases associated with cells can be exploited to provide a means to concentrate therapeutic agents, such as cytotoxic agents, near such cells by providing conjugates that are activated upon cleavage by such enzymes. Such conjugates, upon the action of a cell surface protease or cell-associate protease, release the therapeutic agent, such as a cytotoxic agent, or a derivative thereof that can be converted to a therapeutic agent, locally at the site of action. As noted above, the substrates are designed to be substrates of targeted proteases that are expressed or are active on the surfaces of cells, such as tumor cells or endothelial cells, involved in or present at the site(s) or locus or loci of the disease. By virtue of specific expression, localization or activation of such proteases or the presence of receptors, substrates or enzyme co-factors therefor, administration of the conjugates provided herein permits targeting of therapeutic agents to such cells. Upon contacting with the proteases, active

-10-

therapeutic agents are released in the immediate vicinity of the targeted cells. For example, specific profiles of some of the MTSPs are as follows.

As used herein, "transmembrane serine protease (MTSP)" refers to a family of transmembrane serine proteases that share common structural features as described herein (see, also Hooper *et al.* (2001) *J. Biol. Chem.* 276:857-860). Thus, reference, for example, to "MTSP" encompasses all proteins encoded by the MTSP genes, including but are not limited to: MTSP1, MTSP3, MTSP4, MTSP6, MTSP7, MTSP9, MTSP10, MTSP12, MTSP20, MTSP22 and MTSP25 or an equivalent molecule obtained from any other source or that has been prepared synthetically or that exhibits the same activity. Other MTSPs include, but are not limited to, corin, enteropeptidase, human airway trypsin-like protease (HAT), TMPRSS2 and TMPRSS4. The MTSPs described herein can be used to identify other MTSPs. Methods for isolating nucleic acid encoding other MTSPs, including nucleic acid molecules encoding full-length molecules and splice variants and MTSPs from species, such as cows, sheep, goats, pigs, horses, primates, including chimpanzees and gorillas, rodents, dogs, cats and other species of interest, such as domesticated animals, farm and zoo animals are known to those of skill in the art and are outlined herein. The nucleic acid molecules described herein including those set forth in SEQ IDs can be used to obtain nucleic acid molecules encoding full-length MTSP polypeptides from human sources or from other species, such as by screening appropriate libraries using the nucleic acid molecules or selected primers or probes based thereon.

Sequences of encoding nucleic acid molecules and the encoded amino acid sequences of exemplary MTSPs and/or domains thereof are set forth in SEQ ID Nos. 1-45, 269-270 and 272-276. The term also encompasses MTSPs with amino acid substitutions that do not substantially alter activity of each member and also encompasses polypeptides encoded by splice variants thereof. Hence, encompassed are MTSPs with amino acid substitutions such that the resulting polypeptide retains at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of the proteolytic activity of the unaltered polypeptide, and also encompasses MTSPs encoded by splice variants thereof and MTSPs encoded by allelic variants, such as single nucleotide polymorphisms (SNPs). Suitable

-11-

substitutions, including, although not necessarily, conservative substitutions of amino acids, are known to those of skill in this art and can be made without eliminating the biological activity, such as the catalytic activity, of the resulting molecule. MTSPs include those of animal, such as mammalian, including  
5 human, origin.

As used herein, a "protease domain of an MTSP" refers to an extracellular protease domain of an MTSP that exhibits proteolytic activity and shares homology and structural features with the chymotrypsin/trypsin family protease domains. Hence it is at least the minimal portion of the domain that  
10 exhibits proteolytic activity as assessed by standard *in vitro* assays. Contemplated herein are such protease domains and catalytically active portions thereof.

Exemplary MTSP polypeptides, with the protease domains indicated, are set forth in SEQ ID Nos. 1-45, 269-270 and 272-276, and including smaller  
15 portions thereof that retain or exhibit protease activity. The protease domains vary in size and constitution, including insertions and deletions in surface loops. They retain conserved structure, including at least one of the active site triad, primary specificity pocket, oxyanion hole and/or other features of serine protease domains of proteases. Thus, for purposes herein, the protease domain  
20 is a portion of a MTSP, as defined herein, and is homologous to a domain of other MTSPs. MTSPs include, MTSP1, MTSP3, MTSP4, MTSP6, MTSP7, MTSP9, MTSP10, MTSP12, MTSP20, MTSP22 and MTSP25 (see SEQ ID Nos. 1-19, 42-45, 269-270 and 272-276; see, also International PCT application No. WO 02/00860 (see SEQ ID Nos. 38 and 97 therein, which provide an MTSP12  
25 variant); corin (SEQ ID Nos. 28 and 29), enteropeptidase (SEQ ID Nos. 30 and 31) human airway trypsin-like protease (HAT) (SEQ ID Nos. 32 and 33), hepsin (SEQ ID Nos. 34 and 35), TMPRSS2 (SEQ ID Nos. 36 and 37) and TMPRSS4 (SEQ ID Nos. 38 and 39). As with the larger class of enzymes of the chymotrypsin (S1) fold (see, *e.g.*, Internet accessible MEROPS data base), the  
30 MTSPs protease domains share a high degree of amino acid sequence identity. The His, Asp and Ser residues necessary for activity are present in conserved motifs. In those that are activated by cleavage, the activation site, which



-12-

results in the N-terminus of second chain in the two chain forms has a conserved motif and readily can be identified (see, *e.g.*, amino acids 801-806, SEQ ID No. 29, amino acids 406-410, SEQ ID No. 31; amino acids 186-190, SEQ ID No. 33; amino acids 161-166, SEQ ID No. 35; amino acids 255-259, SEQ ID No. 37; amino acids 190-194, SEQ ID No. 39 and other as known to those of skill and the art and/or as described herein).

For example, with reference to MTSP10 (see SEQ ID Nos. 44 and 45), there are disulfide bonds as follows: C<sub>488</sub>-C<sub>504</sub>, C<sub>587</sub>-C<sub>653</sub>; C<sub>619</sub>-C<sub>632</sub>; C<sub>643</sub>-C<sub>673</sub> (see SEQ ID Nos. 44 and 45) (chymotrypsin numbering 42 to 58; 136-201; 168-182 and 191-220). Disulfide bonds form between the Cys residues C<sub>573</sub>-C<sub>296</sub> to link the protease domain to another domain so that upon activation cleavage (between residues R<sub>462</sub> and I<sub>463</sub> of SEQ ID No. 45) the resulting polypeptide is a two chain molecule. The C<sub>573</sub> (SEQ ID NO. 45) is a free Cys in a single chain form of the protease domain. As noted the protease also can be provided as a two chain molecule. Single chain and two chain forms are proteolytically active. A two chain form is produced by bonding, typically between the C<sub>573</sub> and a Cys outside the protease domain, such as Cys<sub>296</sub>. Upon activation cleavage the disulfide bond remains resulting in a two chain polypeptide. The size of chain "A" is a function of the starting length of the polypeptide prior to activation cleavage between the R<sub>462</sub> and I<sub>463</sub>. Any length polypeptide that includes the protease domain (residues 463-692 of SEQ ID No. 45) or catalytically active fragments thereof, is contemplated herein. Two chain forms include at least the protease domain a polypeptide from C<sub>296</sub> up to and including C<sub>573</sub>.

As used herein, a two-chain form of the protease domain refers to a two-chain form that is formed from a single chain form of the protease in which the Cys pairing between, *e.g.*, a Cys outside the protease domain such as, for example Cys<sub>573</sub> (SEQ ID No. 45 for MTSP), which links the protease domain to the remainder of the polypeptide, the "A" chain. A two chain protease domain form refers to any form in which the "remainder of the polypeptide", *i.e.*, "A" chain, is shortened and includes a Cys from outside the protease domain.

As used herein, the catalytically active domain of an MTSP refers to the protease domain. Reference to the protease domain of an MTSP generally refers

-13-

to a single chain form of the protein. If the two-chain form or both forms is intended, it is so-specified. The zymogen form of each protein is a single chain, which is converted to the active two or multi chain form by activation cleavage. By active form is meant a form active *in vivo* or *in vitro*.

5           As used herein, activation cleavage refers to the cleavage of the protease at the N-terminus of the protease domain (generally between an R and I or V in the full-length protein. By virtue of the Cys-Cys pairing between a Cys outside the protease domain and a Cys in the protease domain (see, *e.g.*, Cys<sub>573</sub> SEQ ID No. 45, upon cleavage the resulting polypeptide has two chains ("A" chain and  
10 the "B" chain, which is the protease domain of an MTSP). Cleavage can be effected by another protease or autocatalytically. The conjugates provided herein advantageously contain sites that are recognized by the active cell surface protease (or cell-associated protease) and are cleaved thereby to release active or an inactive prodrug form of a therapeutic agent.

15           As used herein an MTSP1, whenever referenced herein, includes at least one or all of or any combination of:

                  a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 1 or 40;

                  a polypeptide encoded by a sequence of nucleotides that  
20 hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 1 or 40;

                  a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 2 or 41;

                  a polypeptide that comprises a sequence of amino acids having at  
25 least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 2 or 41; and/or

                  a polypeptide encoded by a splice variant of the MTSP1 set forth  
30 in SEQ ID No. 1 or 40.

          The MTSP1 can be from any animal, particularly a mammal, and includes but is not limited to, humans, rodents, fowl, ruminants and other animals. The

-14-

full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form. MTSP1 also is referred to TADG-15 and matriptase. As described below, the protein originally designated matriptase  
5 appears to be an MTSP1 splice variant or processed product.

As used herein an MTSP3, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 3;

10 a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 3;

a polypeptide that comprises the sequence of amino acids set forth as amino acids 205-437 of SEQ ID No. 4;

15 a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 4; and/or

20 a polypeptide encoded by a splice variant of the MTSP3 set forth in SEQ ID Nos. 3 and 4.

The MTSP3 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain  
25 thereof, including the protease domain, which can be a two chain activated form, or a single chain form.

As used herein an MTSP4, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in  
30 any of SEQ ID No. 5, 7 or 9;

-15-

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in any of SEQ ID Nos. 5, 7 or 9;

a polypeptide that comprises the sequence of amino acids set forth in any of SEQ ID Nos. 6, 8 or 10;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 6, 8 or 10; and/or

a polypeptide encoded by a splice variant of the MTSP4s set forth in SEQ ID Nos. 7-10.

The MTSP4 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.

As used herein an MTSP6, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in any of SEQ ID No. 11;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in any of SEQ ID Nos. 11;

a polypeptide that comprises the sequence of amino acids set forth in any of SEQ ID No. 12;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 12; and/or

-16-

a polypeptide encoded by a splice variant of the MTSP6 set forth in SEQ ID No. 12.

The MTSP6 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full  
5 length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form. Of particular interest herein is the MTSP6 of SEQ ID No. 12.

As used herein an MTSP7, whenever referenced herein, includes at least  
10 one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 13;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence  
15 of nucleotides set forth in SEQ ID No. 13;

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 13;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,  
20 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 14; and/or

a polypeptide encoded by a splice variant of the MTSP7 set forth in SEQ ID No. 13.

25 The MTSP7 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.

30 As used herein an MTSP9, whenever referenced herein, includes at least one or all of or any combination of:

-17-

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 17 or SEQ ID No. 42;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence  
5 of nucleotides set forth in SEQ ID No. 17 or 42;

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 18 or 43;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,  
10 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 18 or 270; and/or

a polypeptide encoded by a splice variant of the MTSP9 set forth in SEQ ID No. 17.

15 The MTSP9 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.

20 As used herein an MTSP10, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 44;

a polypeptide encoded by a sequence of nucleotides that  
25 hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 44;

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 45;

a polypeptide that comprises a sequence of amino acids having at  
30 least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or

-18-

99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 45; and/or

a polypeptide encoded by a splice variant of the MTSP10 set forth in SEQ ID No. 44.

- 5           The MTSP10 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.
- 10           MTSP10 polypeptides, including, but not limited to splice variants thereof, and nucleic acids encoding MTSPs, and domains, derivatives and analogs thereof are provided herein. Single chain protease domains that have an N-terminus functionally equivalent to that generated by activation of the zymogen form of MTSP10 are also provided. The cleavage site for the protease
- 15   domain of MTSP10 is between amino acid R and amino acids I (R↓IIGGT) (residues 462-467 SEQ ID No. 45).

As used herein an MTSP12, whenever referenced herein, includes at least one or all of or any combination of: SEQ ID No. 19 and 20

- a polypeptide encoded by the sequence of nucleotides set forth in
- 20   SEQ ID No. 19 or by a sequence of nucleotides that includes nucleotides that encode the sequence of amino acids set forth in SEQ ID No. 20;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in is set forth as SEQ ID No. 19;

- 25           a polypeptide that includes the sequence of amino acids set forth in SEQ ID No. 20 or a catalytically active portion thereof;

- a polypeptide that includes a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or
- 30   99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 20; and/or

-19-

a polypeptide encoded by a splice variant of the MTSP12 that includes the sequence of amino acids set forth in SEQ ID No. 20.

In particular, the MTSP12 polypeptide, with the protease domains as indicated in SEQ ID Nos. 19 and 20, is provided. The polypeptide is a single or  
5 multi-chain polypeptide. A protease domain of an MTSP12, whenever referenced herein, includes at least one or all of or any combination of or a catalytically active portion of:

a polypeptide that includes the sequence of amino acids set forth in SEQ ID No. 20 or a catalytically active portion thereof but that does not  
10 include the sequence of amino acids set forth in SEQ ID No. 271;

a polypeptide that includes the sequence of amino acids set forth in SEQ ID No. 272 or a catalytically active fragment thereof;

a polypeptide containing amino acids 237 to 456 of SEQ ID No. 6, a polypeptide containing amino acids 538 to 765 of SEQ ID No. 20, and a  
15 polypeptide containing amino acids 861 to 1087 of SEQ ID No. 20, but that does not include the sequence of amino acids set forth in SEQ ID No. 271;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to a sequence of nucleotides that encodes any of the polypeptides of a)-c);

20 a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 20 but that does not encode the sequence of amino acids set forth in SEQ ID No. 271;

a polypeptide that includes a sequence of amino acids having at  
25 least about 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 20;

a polypeptide that includes a sequence of amino acids having at  
least about 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,  
30 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids of the polypeptides of a)-e);



-20-

a polypeptide encoded by a splice variant of a sequence of nucleotides that encodes an MTSP12 of any of the above.

Smaller portions thereof that retain protease activity are also provided. The MTSP12 can be from any animal, particularly a mammal, and includes but  
5 are not limited to, humans, rodents, fowl, ruminants and other animals. The full-length zymogen or two-chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two-chain activated form, or a single chain form. MTSP12 also includes the variant described International PCT application No. WO 02/00860 (see SEQ ID Nos. 38 and 97  
10 therein).

As used herein an MTSP20, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 273;

15 a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 273;

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 273;

20 a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 274; and/or

25 a polypeptide encoded by a splice variant of the MTSP20 encoded by the sequence of nucleotides that includes the sequence set forth in SEQ ID No. 273.

The MTSP20 may be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other  
30 animals. The full length zymogen or two-chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two-chain activated form, or a single chain form.

As used herein an MTSP22, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 275;

- 5 a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 275;

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 276;

- 10 a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 276; and/or

- 15 a polypeptide encoded by a splice variant of the MTSP22 set forth in SEQ ID No. 275.

The MTSP22 may be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two-chain activated form is contemplated  
20 or any domain thereof, including the protease domain, which can be a two-chain activated form, or a single chain form.

As used herein an MTSP25, whenever referenced herein, includes at least one or all of or any combination of:

- a polypeptide encoded by the sequence of nucleotides set forth in  
25 SEQ ID No. 269;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 269;

- a polypeptide that comprises the sequence of amino acids set  
30 forth in SEQ ID No. 270;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,

-22-

87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 270; and/or

a polypeptide encoded by a splice variant of the MTSP25 set forth  
5 in SEQ ID No. 269.

The MTSP25 may be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two-chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two-chain  
10 activated form, or a single chain form.

As used herein, a human protein is one encoded by nucleic acid present in the genome of a human, including all allelic variants and conservative variations as long as they are not variants found in other mammals.

As used herein, not substantially cleaved by plasmin or prostate specific  
15 antigen (PSA) (or other non-cell surface-associated protease), means in comparable *in vitro* assays (under optimal conditions for each enzyme) in which the activity of a targeted cell surface membrane protease or catalytically active portion of the activity of the protease domain (or a catalytically active form thereof) of prostate specific antigen (PSA) or plasmin for cleavage of the  
20 conjugate is compared, the relative activity is greater than at least 2:1, 3:1, 4:1, 5:1, 10:1, 50:1 or 100:1.

As used herein, activity refers to the ratio  $k_{cat}/K_m$ , where  $k_{cat}$  is the rate of catalytic turnover for a particular enzyme, and  $K_m$  is the Michaelis constant for the binding of the substrate.

25 As used herein, a "nucleic acid encoding a protease domain or catalytically active portion of a MTSP" shall be construed as referring to a nucleic acid encoding only the recited single chain protease domain or active portion thereof, and not the other contiguous portions of the MTSP as a continuous sequence.

30 As used herein, a CUB domain is a motif that mediates protein-protein interactions in complement components

-23-

C1r/C1s and has also been identified in various proteins involved in developmental processes.

As used herein, a zymogen is an enzymatically inactive protein (i.e., typically, but not necessarily, less than 1% of active form) that is converted to a proteolytic enzyme by the action of an activator, including by autoactivation. Inactive means less active than the form those of skill in the art consider to be the active form of the enzyme. The ratio of activity of a zymogen to the activated form varies from enzyme-to-enzyme.

As used herein, "disease or disorder" refers to a pathological condition in an organism resulting from, e.g., infection or genetic defect, and characterized by identifiable symptoms. The diseases contemplated for treatment herein are any for which a cell surface protease, including a cell-localized or cell-associated protease is associated with a targeted cell or tissue involved in the disease or disease process. Such association can be because the protease is involved in the disease or is serendipitously associated with cells involved with the disease. These diseases herein are called cell surface protease-associated diseases. Hence, to treat the disease a cell surface protease is identified that is expressed on cells associated with the disorder, such as, for example, immune cells for treating inflammatory diseases, and virally infected cells for treating viral diseases. The conjugate is designed as described herein for cleavage by the selected protease.

As used herein, neoplasm (neoplasia) refers to abnormal new growth, and thus means the same as *tumor*, which can be benign or malignant. Unlike *hyperplasia*, neoplastic proliferation persists even in the absence of the original stimulus.

As used herein, neoplastic disease refers to any disorder involving cancer, including tumor development, growth, metastasis and progression.

As used herein, cancer refers to a general term for diseases caused by any type of malignant tumor.

As used herein, malignant, as applied to tumors, refers to primary tumors that have the capacity of *metastasis* with loss of *growth control* and *positional control*.

-24-

As used herein, endotheliase refers to a mammalian protein, including human protein, that has a transmembrane domain and is expressed or active on the surface of endothelial cells and includes a protease domain, particularly an extracellular protease domain, and is generally a serine protease (see, also U.S. application Serial No. 09/717,473 and International PCT application No. WO 01/36604). Thus, reference, for example, to endotheliase encompasses all proteins encoded by the endotheliase gene family, or an equivalent molecule obtained from any other source or that has been prepared synthetically or that exhibits the same activity. The endotheliase gene family are transmembrane proteases expressed or active in endothelial cells. These proteases include serine proteases. These include proteins that have these features and also include a protease domain that exhibits sequence homology to the endothelias 1 and 2. Endotheliase 1 and 2, for example exhibit about 40% or 45% identity. Sequence homology means sequence identity along its length when aligned to maximize identity of at least about 25%, 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater number of residues. Sequence homology also is assessed by determining whether the encoding sequences of nucleic acids hybridize under conditions of at least moderate, or for more closely related proteins, high stringency to the nucleic acid molecules provided herein or to those that encode the same proteins but differ in sequence by virtue of the degeneracy of the genetic code. In addition, "endothelias" encompasses endothelias with amino acid substitutions, including those set forth in Table 1, such that the resulting polypeptide retains at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of the proteolytic activity of the unaltered polypeptide. Suitable substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity of the resulting molecule. As noted, those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Bejacmin/Cummings Pub. Co., p.224). Also

-25-

included within the definition of "endotheliases", is the catalytically active fragment or shed forms of an endotheliase.

As used herein an endotheliase 1, whenever referenced herein, includes at least one or all of or any combination of:

- 5                   a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 21;  
                  a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 21;
- 10                  a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 22;  
                  a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or
- 15                  99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 22; and/or  
                  a polypeptide encoded by a splice variant of a nucleic acid molecule that encodes a protein containing the polypeptide set forth in SEQ ID No. 22.
- 20                  The endotheliase 1 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.
- 25                  As used herein an endotheliase 2, whenever referenced herein, includes at least one or all of or any combination of:  
                  a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 23 or 25;  
                  a polypeptide encoded by a sequence of nucleotides that
- 30                  hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 23 or 25;

-26-

a polypeptide that comprises the sequence of amino acids set forth in SEQ ID No. 24 or 26;

a polypeptide that comprises a sequence of amino acids having at least about 40%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,  
5 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 24 or 26; and/or

a polypeptide encoded by a splice variant of a nucleic acid set forth in SEQ ID No. 23 or 25.

10 The endotheliase 2 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full length zymogen or two chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two chain activated form, or a single chain form.

15 As used herein, the protease domain of an endotheliase refers to the polypeptide portion of the endotheliase that is the extracellular portion that exhibits protease activity. The protease domain is a polypeptide that includes at least the minimum number of amino acids, generally more than 50 or 100, required for protease activity. Protease activity can be assessed empirically,  
20 such as by testing the polypeptide for its ability to act as a protease. Assays, such as those described in the EXAMPLES, with the exception that a known endotheliase substrate is employed in place of the test compounds, can be used to assess protease activity. Furthermore, since proteases, particularly serine proteases, have characteristic structures and sequences or motifs, the protease  
25 domain can be readily identified by such structure and sequence or motif.

As used herein, a portion of protease domain of endotheliase refers to a portion of protease domain of endotheliase that is located within or is the extracellular domain of an endotheliase and exhibits serine proteolytic activity. Hence, it is at least the minimal portion of the extracellular domain that exhibits  
30 proteolytic activity as assessed by standard assays. An exemplary protease domain of an endotheliase is set forth in SEQ ID No. 22 and as amino acids 321-688 and 321-562 of SEQ ID Nos. 24 and 26, respectively. Smaller portions

-27-

thereof that retain protease activity are contemplated. The protease domains vary in size and constitution, including insertions and deletions in surface loops. Such domains exhibit conserved structure, including at least one structural feature, such as the active site triad, primary specificity pocket, oxyanion hole and/or other features of serine protease domains of proteases. Thus, for purposes herein, the protease domain is a portion of an endotheliase, as defined herein, but is homologous in terms of structural features and retention of sequence of similarity or homology the protease domain of chymotrypsin or trypsin.

10 As used herein, homologous means about greater than about 25%, 40%, 60%, 80%, 90%, 95%, 98% or greater sequence identity. By sequence identity, the number of conserved amino acids as determined by standard alignment algorithms programs, and used with default gap penalties established by each supplier. Also homology can be assessed by conserved nucleic acid  
15 sequence, which includes anything that hybridizes under at least low stringency conditions and encodes the domain. Similarly, nucleic acid sequence alignment programs are commercially available (DNASTar "MegAlign" program (Madison, WI) and the University of Wisconsin Genetics Computer Group (UWG) "Gap" program (Madison, WI)). Substantially homologous nucleic acid molecules  
20 would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid of interest. Also contemplated are nucleic acid molecules that contain degenerate codons in place of codons in the hybridizing nucleic acid molecule.

As used herein, recitation that a polypeptide consists essentially of the  
25 protease domain means that the only endotheliase portion of the polypeptide is a protease domain or a catalytically active portion thereof. The polypeptide can optionally include additional non-endotheliase-derived sequences of amino acids.

As used herein, domain refers to a portion of a molecule, *e.g.*, proteins or nucleic acids, that is structurally and/or functionally distinct from other  
30 portions of the molecule.



-28-

As used herein, an active form of a protease refers to an enzyme that catalyzes hydrolysis of proteins or peptides. Reference to a protease includes the active and zymogen or other less active form.

As used herein, nucleic acids include DNA, RNA and analogs thereof,  
5 including peptide nucleic acids (PNA) and mixtures thereof. Nucleic acids can be single or two stranded. When referring to probes or primers, optionally labeled, with a detectable label, such as a fluorescent or radiolabel, single-stranded molecules are contemplated. Such molecules are typically of a length such that their targets are statistically unique or of low copy number (typically less than 5,  
10 generally less than 3) for probing or priming a library. Generally a probe or primer contains at least 14, 16 or 30 contiguous of sequence complementary to or identical to a gene of interest. Probes and primers can be 10, 20, 30, 50, 100 or more nucleic acids long.

As used herein, nucleic acid encoding a fragment or portion of an  
15 endotheliase refers to a nucleic acid encoding only the recited fragment or portion of endotheliase protein, and not the other contiguous portions of the endotheliase as a continuous sequence.

As used herein, heterologous nucleic acid is nucleic acid that, if it is DNA encodes RNA, or, if RNA, encodes proteins that generally are not normally  
20 produced *in vivo* by the cell in which it is expressed or that mediates or encodes mediators that alter expression of endogenous nucleic acid, such as DNA, by affecting transcription, translation, or other regulatable biochemical processes or that is located in a different locus from its normal locus. Heterologous nucleic acid is generally not endogenous to the cell into which it is introduced, but has  
25 been obtained from another cell or prepared synthetically. Generally, although not necessarily, such nucleic acid encodes RNA and proteins that are not normally produced by the cell in which it is now expressed.

Heterologous nucleic acid, such as DNA, also be referred to as foreign nucleic acid, such as DNA. Any nucleic acid, such as DNA, that one of skill in  
30 the art would recognize or consider as heterologous or foreign to the cell in which is expressed is herein encompassed by heterologous nucleic acid; heterologous nucleic acid includes exogenously added nucleic acid that is also

-29-

expressed endogenously. Examples of heterologous nucleic acid include, but are not limited to, nucleic acid that encodes traceable marker proteins, such as a protein that confers drug resistance, nucleic acid that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and nucleic acid, such as DNA, that encodes other types of proteins, such as antibodies, and RNA, such as RNA interference (RNAi) or other double-stranded RNA, and antisense RNA. Antibodies that are encoded by heterologous nucleic acid can be secreted or expressed on the surface of the cell in which the heterologous nucleic acid has been introduced.

- 10 For example, nucleic acid can be the the targeted agent, such as the therapeutic or diagnostic agent, in the conjugate. Nucleic acids, include ds RNA use for RNA interference (RNAi) (see, *e.g.* Chuang *et al.* (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:4985) which is employed to inhibit the expression of a targeted gene by generating loss-of-function. Methods relating to the use of
- 15 RNAi to silence genes in organisms including, mammals, *C. elegans*, *Drosophila* and plants, and humans are known (see, *e.g.*, Fire *et al.* (1998) *Nature* 391:806-811; Fire (1999) *Trends Genet.* 15:358-363; Sharp (2001) *Genes Dev.* 15:485-490; Hammond, *et al.* (2001) *Nature Rev. Genet.* 2:110-1119; Tuschl (2001) *Chem. Biochem.* 2:239-245; Hamilton *et al.* (1999) *Science* 286:950-952;
- 20 Hammond *et al.* (2000) *Nature* 404:293-296; Zamore *et al.* (2000) *Cell* 101:25-33; Bernstein *et al.* (2001) *Nature* 409: 363-366; Elbashir *et al.* (2001) *Genes Dev.* 15:188-200; Elbashir *et al.* (2001) *Nature* 411:494-498; International PCT application No. WO 01/29058; International PCT application No. WO 99/32619). By selecting appropriate sequences, expression of dsRNA can
- 25 interfere with accumulation of endogenous mRNA encoding a targeted gene product. Regions that include at least about 21 nucleotides and that are selective (i.e. whose target is unique) for the nucleic acid encoding a targeted gene product are used to prepare the RNAi.

-30-

As used herein, genetic therapy involves the transfer of heterologous nucleic acid, such as DNA, into certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The nucleic acid molecules are included in a conjugate linked via a cell surface protein cleavage site. The nucleic acid, such as DNA, is introduced into the selected target cells in a manner such that the heterologous nucleic acid, such as DNA, is expressed and a therapeutic product encoded thereby is produced. Alternatively the heterologous nucleic acid, such as DNA, can in some manner mediate expression of DNA that encodes the therapeutic product, or it can encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy can also be used to deliver nucleic acid encoding a gene product that replaces a defective gene or supplements a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid can encode a therapeutic compound, such as a growth factor inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous nucleic acid, such as DNA, encoding the therapeutic product can be modified prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof. Genetic therapy can also involve delivery of an inhibitor or repressor or other modulator of gene expression, such dsRNA or antisense or other nucleic acid molecule. The conjugates herein can be used to deliver a product, such as a nucleic acid for gene therapy.

As used herein, a therapeutically effective product for gene therapy is a product that is encoded by heterologous nucleic acid, typically DNA, that, upon introduction of the nucleic acid into a host, a product is expressed that ameliorates or eliminates the symptoms, manifestations of an inherited or acquired disease or that cures the disease. Also included are biologically active nucleic acid molecules, such as RNAi and antisense.

As used herein, a sequence complementary to at least a portion of an RNA, with reference to antisense oligonucleotides, means a sequence having sufficient complementarity to be able to hybridize with the RNA, generally under moderate or high stringency conditions, forming a stable duplex; in the case of double-stranded SP antisense nucleic acids, a single strand of the duplex DNA (or dsRNA) can thus be tested, or triplex formation can be assayed. The ability to hybridize depends on the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with a SP encoding RNA it can contain and still form a stable duplex (or triplex, as the case can be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Amino acid substitutions can be made or occur in any SPs and protease domains thereof. Amino acid substitutions include conservative substitutions, such as those set forth in Table 1, which do not eliminate proteolytic activity. As described herein, substitutions that alter properties of the proteins, such as removal of cleavage sites and other such sites are also contemplated; such substitutions are generally non-conservative, but can be readily effected by those of skill in the art.

Suitable conservative substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity, for example enzymatic activity, of the resulting molecule. Also included within the definition, is the catalytically active fragment of an SP, particularly a single chain protease portion.

Conservative amino acid substitutions are made, for example, in accordance with those set forth in TABLE 1 as follows:

-32-

TABLE 1

	Ala (A)	Gly; Ser
	Arg (R)	Lys, Orn
5	Asn (N)	Gln; His
	Asp (D)	Glu
	Cys (C)	Ser
	Gln (Q)	Asn
	Glu (E)	Asp
10	Gly (G)	Ala; Pro
	His (H)	Asn; Gln
	Ile (I)	Leu; Val; Nle; Met
	Leu (L)	Ile; Val; Nle; Met
	Lys (K)	Arg; Gln; Glu
15	Met (M)	Leu; Tyr; Ile; Nle
	Phe (F)	Met; Leu; Tyr, Trp
	Ser (S)	Thr
	Thr (T)	Ser
	Trp (W)	Tyr; Phe
20	Tyr (Y)	Trp; Phe
	Val (V)	Ile; Leu; Nle; Met

Other substitutions are also permissible and can be determined empirically or in accord with known conservative substitutions. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence can be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their well-known, three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used

-33-

routinely in the art. Other abbreviations, include: hR or hArg for homoarginine; hY or hTyr for homotyrosine; Cha for cyclohexylalanine; Amf for 4-aminomethylphenylalanine; DPL for 2-(4,6-dimethylpyrimidinyl)lysine; (imidazolyl)K for N'-(2-imidazolyl)lysine; Me2PO3-Y for

- 5 O-dimethylphosphotyrosine; O-Me-Y for O-methyltyrosine; TIC for tetrahydro-3-isoquinoline carboxylic acid; MeL for 2-keto-3-amino-5-methyl-hexane; DAP for 1,3-diaminopropane; TFA for trifluoroacetic acid; AA for acetic acid.

- As used herein, a splice variant refers to a variant produced by  
10 differential processing of a primary transcript of genomic DNA that results in more than one type of mRNA.

As used herein, a probe or primer based on a nucleotide sequence disclosed herein, includes at least 10, 14, generally at least 16 or 30 or 100 contiguous sequence of nucleotides.

- 15 As used herein, antisense polynucleotides refer to synthetic sequences of nucleotide bases complementary to mRNA or the sense strand of double-stranded DNA. Admixture of sense and antisense polynucleotides under appropriate conditions leads to the binding of the two molecules, or hybridization. When these polynucleotides bind to (hybridize with) mRNA,  
20 inhibition of protein synthesis (translation) occurs. When these polynucleotides bind to double-stranded DNA, inhibition of RNA synthesis (transcription) occurs. The resulting inhibition of translation and/or transcription leads to an inhibition of the synthesis of the protein encoded by the sense strand. Antisense nucleic acid molecules typically contain a sufficient number of nucleotides to specifically  
25 bind to a target nucleic acid, generally at least 5 contiguous nucleotides, often at least 14 or 16 or 30 contiguous nucleotides or modified nucleotides complementary to the coding portion of a nucleic acid molecule that encodes a gene of interest, for example, nucleic acid encoding a single chain protease domain of an SP.

- 30 As used herein, an array refers to a collection of elements, such as antibodies, containing three or more members. An addressable array is one in which the members of the array are identifiable, typically by position on a solid

-34-

phase support. Hence, in general the members of the array are immobilized on discrete identifiable loci on the surface of a solid phase.

As used herein, antibody refers to an immunoglobulin, whether natural or partially or wholly synthetically produced, including any derivative thereof that  
5 retains the specific binding ability of the antibody. Hence antibody includes any protein having a binding domain that is homologous or substantially homologous to an immunoglobulin binding domain. Antibodies include members of any immunoglobulin claims, including IgG, IgM, IgA, IgD and IgE.

As used herein, antibody fragment refers to any derivative of an antibody  
10 that is less than full-length, retaining at least a portion of the full-length antibody's specific binding ability. Examples of antibody fragments include, but are not limited to, Fab, Fab', F(ab)<sub>2</sub>, single-chain Fvs (scFV), FV, dsFV diabody and Fd fragments. The fragment can include multiple chains linked together, such as by disulfide bridges. An antibody fragment generally contains at least  
15 about 50 amino acids and typically at least 200 amino acids.

As used herein, an Fv antibody fragment is composed of one variable heavy domain (V<sub>H</sub>) and one variable light domain linked by noncovalent interactions.

As used herein, a dsFV refers to an Fv with an engineered intermolecular  
20 disulfide bond, which stabilizes the V<sub>H</sub>-V<sub>L</sub> pair.

As used herein, an F(ab)<sub>2</sub> fragment is an antibody fragment that results from digestion of an immunoglobulin with pepsin at pH 4.0-4.5; it can be recombinantly expressed to produce the equivalent fragment.

As used herein, Fab fragments are antibody fragments that result from  
25 digestion of an immunoglobulin with papain; they can be recombinantly expressed to produce the equivalent fragment.

As used herein, scFVs refer to antibody fragments that contain a variable light chain (V<sub>L</sub>) and variable heavy chain (V<sub>H</sub>) covalently connected by a polypeptide linker in any order. The linker is of a length such that the two  
30 variable domains are bridged without substantial interference. Exemplary linkers include, but are not limited to, (Gly-Ser)<sub>n</sub> residues, which can include one Glu or Lys residues dispersed throughout, for example, to increase solubility.

-35-

As used herein, humanized antibodies refer to antibodies that are modified to include human sequences of amino acids so that administration to a human does not provoke an immune response. Methods for preparation of such antibodies are known. For example, to produce such antibodies, the  
5 encoding nucleic acid in the hybridoma or other prokaryotic or eukaryotic cell, such as an *E. coli* or a CHO cell, that expresses the monoclonal antibody is altered by recombinant nucleic acid techniques to express an antibody in which the amino acid composition of the non-variable region is based on human antibodies. Computer programs have been designed to identify such non-  
10 variable regions.

As used herein, diabodies are dimeric scFV; diabodies typically have shorter peptide linkers than scFVs, and they generally dimerize.

As used herein, production by recombinant means by using recombinant DNA methods means the use of the well known methods of molecular biology  
15 for expressing proteins encoded by cloned DNA.

As used herein, the term assessing is intended to include quantitative and qualitative determination in the sense of obtaining an absolute value for the activity of an SP, or a domain thereof, present in the sample, and also of obtaining an index, ratio, percentage, visual or other value indicative of the level  
20 of the activity. Assessment can be direct or indirect and the chemical species actually detected need not of course be the proteolysis product itself but can for example be a derivative thereof or some further substance.

As used herein, biological activity refers to the *in vivo* activities of a compound or physiological responses that result upon *in vivo* administration of a  
25 compound, composition or other mixture. Biological activity, thus, encompasses therapeutic effects and pharmaceutical activity of such compounds, compositions and mixtures. Biological activities can be observed in *in vitro* systems designed to test or use such activities.

As used herein, a combination refers to any association between two or  
30 among more items.



-36-

As used herein, fluid refers to any composition that can flow. Fluids thus encompass compositions that are in the form of semi-solids, pastes, solutions, aqueous mixtures, gels, lotions, creams and other such compositions.

As used herein, an effective amount of a compound for treating a  
5 particular disease is an amount that is sufficient to ameliorate, or in some manner reduce the symptoms associated with the disease. Such amount can be administered as a single dosage or can be administered according to a regimen, whereby it is effective. The amount can cure the disease but, typically, is administered in order to ameliorate the symptoms of the disease. Repeated  
10 administration can be required to achieve the desired amelioration of symptoms.

As used herein, equivalent, when referring to two sequences of nucleic acids, means that the two sequences in question encode the same sequence of amino acids or equivalent proteins. When equivalent is used in referring to two proteins or peptides, it means that the two proteins or peptides have  
15 substantially the same amino acid sequence with amino acid substitutions (see, *e.g.*, Table 1, above) that do not substantially alter the activity or function of the protein or peptide (*i.e.*, retain at least about 1 % of the activity). When equivalent refers to a property, the property does not need to be present to the same extent (*e.g.*, two peptides can exhibit different rates of the same type of  
20 enzymatic activity), but the activities are generally substantially the same. Complementary, when referring to two nucleotide sequences, means that the two sequences of nucleotides are capable of hybridizing, typically with less than 25 %, often with less than 15 %, or even less than 5 % or with no mismatches between opposed nucleotides. Generally the two molecules hybridize under  
25 conditions of high stringency.

As used herein, a method for treating or preventing disease or disorder associated with undesired and/or uncontrolled angiogenesis means that the diseases or the symptoms associated with the undesired and/or uncontrolled angiogenesis are alleviated, reduced, ameliorated, prevented, placed in a state of  
30 remission, or maintained in a state of remission. It also means that the hallmarks of pathological angiogenesis are eliminated, reduced or prevented by the treatment. Non-limiting examples of the hallmarks of the pathological

-37-

angiogenesis include uncontrolled degradation of the basement membrane and proximal extracellular matrix of the endothelial cells, migration, division, and organization of the endothelial cells into new functioning capillaries, and the persistence of such functioning capillaries.

5           As used herein, operatively linked or operationally associated refers to the functional relationship of DNA with regulatory and effector sequences of nucleotides, such as promoters, enhancers, transcriptional and translational stop sites, and other signal sequences. For example, operative linkage of DNA to a promoter refers to the physical and functional relationship between the DNA and  
10 the promoter such that the transcription of such DNA is initiated from the promoter by an RNA polymerase that specifically recognizes, binds to and transcribes the DNA. In order to optimize expression and/or *in vitro* transcription, it can be necessary to remove, add or alter 5' untranslated portions of the clones to eliminate extra, potential inappropriate alternative  
15 translation initiation (*i.e.*, start) codons or other sequences that can interfere with or reduce expression, either at the level of transcription or translation. Alternatively, consensus ribosome binding sites (see, *e.g.*, Kozak (1991) *J. Biol. Chem.* 266:19867-19870) can be inserted immediately 5' of the start codon and can enhance expression. The desirability of (or need for) such modification  
20 can be empirically determined.

          As used herein, a promoter region or promoter element refers to a segment of DNA or RNA that controls transcription of the DNA or RNA to which it is operatively linked. The promoter region includes specific sequences that are sufficient for RNA polymerase recognition, binding and transcription initiation.  
25 This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, binding and transcription initiation activity of RNA polymerase. These sequences can be *cis* acting or can be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, can be constitutive or regulated. Exemplary  
30 promoters contemplated for use in prokaryotes include the bacteriophage T7 and T3 promoters.

-38-

As used herein, sample refers to anything which can contain an analyte for which an analyte assay is desired. The sample can be a biological sample, such as a biological fluid or a biological tissue. Examples of biological fluids include urine, blood, plasma, serum, saliva, semen, stool, sputum, cerebral  
5 spinal fluid, tears, mucus, amniotic fluid or the like. Biological tissues are aggregates of cells, usually of a particular kind together with their intercellular substance that form one of the structural materials of a human, animal, plant, bacterial, fungal or viral structure, including connective, epithelium, muscle and nerve tissues. Examples of biological tissues also include organs, tumors, lymph  
10 nodes, arteries and individual cell(s).

As used herein, to hybridize under conditions of a specified stringency is used to describe the stability of hybrids formed between two single-stranded DNA fragments and refers to the conditions of ionic strength and temperature at which such hybrids are washed, following annealing under conditions of  
15 stringency less than or equal to that of the washing step. Typically high, medium and low stringency encompass the following conditions or equivalent conditions thereto:

- 1) high stringency: 0.1 x SSPE or SSC, 0.1% SDS, 65°C
- 2) medium stringency: 0.2 x SSPE or SSC, 0.1% SDS, 50°C
- 20 3) low stringency: 1.0 x SSPE or SSC, 0.1% SDS, 50°C.

Equivalent conditions refer to conditions that select for substantially the same percentage of mismatch in the resulting hybrids. Additions of ingredients, such as formamide, Ficoll, and Denhardt's solution affect parameters such as the temperature under which the hybridization should be conducted and the rate of  
25 the reaction. Thus, hybridization in 5 X SSC, in 20% formamide at 42° C is substantially the same as the conditions recited above hybridization under conditions of low stringency. The recipes for SSPE, SSC and Denhardt's and the preparation of deionized formamide are described, for example, in Sambrook *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor  
30 Laboratory Press, Chapter 8; see, Sambrook *et al.*, vol. 3, p. B.13, see, also, numerous catalogs that describe commonly used laboratory solutions). It is

-39-

understood that equivalent stringencies can be achieved using alternative buffers, salts and temperatures.

The terms substantially identical or similar varies with the context as understood by those skilled in the relevant art and generally means at least 40,  
5 60, 80, 90, 95 or 98%.

As used herein, substantially identical to a product means sufficiently similar so that the property of interest is sufficiently unchanged so that the substantially identical product can be used in place of the product.

As used herein, target cell refers to a cell that expresses a cell surface  
10 protease.

As used herein, test substance, including compounds provided herein, refers to a chemically defined compound (*e.g.*, organic molecules, inorganic molecules, organic/inorganic molecules, proteins, peptides, nucleic acids, oligonucleotides, lipids, polysaccharides, saccharides, or hybrids among these  
15 molecules such as glycoproteins, etc.) or mixtures of compounds (*e.g.*, a library of test compounds, natural extracts or culture supernatants, etc.) whose effect on or interaction with a cell surface protein or cell surface-associated protein, or a domain thereof, is determined by the methods herein.

As used herein, the terms a therapeutic agent, therapeutic regimen,  
20 radioprotectant, chemotherapeutic mean conventional drugs and drug therapies, including vaccines, which are known to those skilled in the art. Radiotherapeutic agents are well known in the art.

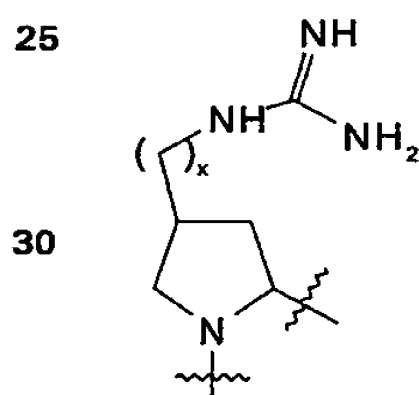
As used herein, vector (or plasmid) refers to discrete elements that are used to introduce heterologous DNA into cells for expression and/or replication  
25 thereof. The vectors typically remain episomal, but can be designed to effect integration of a gene or portion thereof into a chromosome of the genome. Also contemplated are vectors that are artificial chromosomes, such as yeast artificial chromosomes and mammalian artificial chromosomes. Selection and use of such vehicles are well known to those of skill in the art. An expression vector  
30 includes vectors capable of expressing DNA that is operatively linked with regulatory sequences, such as promoter regions, that are capable of effecting expression of such DNA fragments. Thus, an expression vector refers to a

-40-

recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector that, upon introduction into an appropriate host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome.

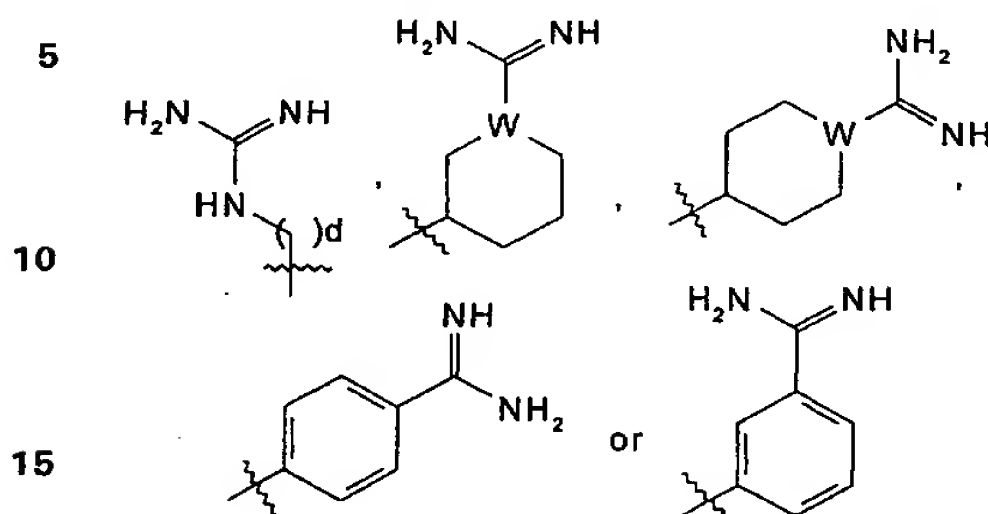
As used herein, chemically stable means that the compound is stable enough to be formulated for pharmaceutical use. Such chemical stability is well known to those of skill in the art and can be determined by well known routine methods. Whether a given compound is chemically stable enough to be formulated for pharmaceutical use depends on a number of factors including, but not limited to, the type of formulation and route of administration desired, the disease to be treated, and the method of preparing the pharmaceutical formulation.

As used herein, a "functional equivalent" of a side chain of an amino acid is a group or moiety that functions in substantially the same way as the naturally occurring side chain to achieve substantially the same result (*e.g.*, a substrate for a cell surface protease). For example, functional equivalents of the side chain of arginine include, but are not limited to, homoarginine, guanidinoaminopropyl, guanidinoaminoethyl, (Me)<sub>2</sub>arginine side chain, (Et)<sub>2</sub>arginine side chain, (4-aminomethyl)phenylmethyl, 4-amidinophenylmethyl, 4-guanidinophenylmethyl, or a conformationally constrained arginine side chain analog such as:



-41-

where x is 0 or 1 (see, e.g., Webb *et al.* (1991) *J. Org. Chem.* 56:3009), or a conformationally constrained arginine side chain analog such as:



where d is an integer from 0 to 5, or 1 to 3; and W is N or CH; or a mono- or di-substituted N-alkyl derivative of the above groups, where alkyl is, in certain embodiments, lower alkyl, such as methyl.

As used herein, pharmaceutically acceptable derivatives of a compound include salts, esters, enol ethers, enol esters, acids, bases, solvates, hydrates or prodrugs thereof. Such derivatives can be readily prepared by those of skill in this art using known methods for such derivatization. The compounds produced can be administered to animals or humans without substantial toxic effects and either are pharmaceutically active or are prodrugs. Pharmaceutically acceptable salts include, but are not limited to, amine salts, such as but not limited to N,N'-dibenzylethylenediamine, chlorprocaine, choline, ammonia, diethanolamine and other hydroxyalkylamines, ethylenediamine, N-methylglucamine, procaine, N-benzylphenethylamine, 1-para-chlorobenzyl-2-pyrrolidin-1'-ylmethylbenzimidazole, diethylamine and other alkylamines, piperazine and tris(hydroxymethyl)aminomethane; alkali metal salts, such as but not limited to lithium, potassium and sodium; alkali earth metal salts, such as but not limited to barium, calcium and magnesium; transition metal salts, such as but not limited to zinc; and other metal salts, such as but not limited to sodium hydrogen phosphate and disodium phosphate; and also including, but not limited to, salts

of mineral acids, such as but not limited to hydrochlorides and sulfates; and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates. Pharmaceutically acceptable esters include, but are not limited to, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl and heterocyclyl esters of acidic groups, including, but not limited to, carboxylic acids, phosphoric acids, phosphinic acids, sulfonic acids, sulfinic acids and boronic acids. Pharmaceutically acceptable enol ethers include, but are not limited to, derivatives of formula  $C=C(OR)$  where R is hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl or heterocyclyl. Pharmaceutically acceptable enol esters include, but are not limited to, derivatives of formula  $C=C(OC(O)R)$  where R is hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl or heterocyclyl. Pharmaceutically acceptable solvates and hydrates are complexes of a compound with one or more solvent or water molecule, generally 1 to about 100, typically 1 to about 10, such as 1 to about 2, 3 or 4, solvent or water molecules.

As used herein, treatment means any manner in which one or more of the symptoms of a condition, disorder or disease are ameliorated or otherwise beneficially altered. Treatment also encompasses any pharmaceutical use of the compositions herein, such as use for treating cancer.

As used herein, amelioration of the symptoms of a particular disorder by administration of a particular pharmaceutical composition refers to any lessening, whether permanent or temporary, lasting or transient that can be attributed to or associated with administration of the composition.

As used herein, a prodrug is a compound that, upon *in vivo* administration, is metabolized or otherwise converted to the biologically, pharmaceutically or therapeutically active form of the compound. To produce a prodrug, the pharmaceutically active compound is modified such that the active compound is regenerated by metabolic processes. The prodrug can be designed to alter the metabolic stability or the transport characteristics of a drug, to mask side effects or toxicity, to improve the flavor of a drug or to alter other

characteristics or properties of a drug. By virtue of knowledge of pharmacodynamic processes and drug metabolism *in vivo*, those of skill in this art, once a pharmaceutically active compound is known, can design prodrugs of the compound (see, *e.g.*, Nogrady (1985) *Medicinal Chemistry A Biochemical Approach*, Oxford University Press, New York, pages 388-392).

It is to be understood that the conjugates provided herein can contain chiral centers. Such chiral centers can be of either the (R) or (S) configuration, or can be a mixture thereof. Thus, the compounds provided herein can be enantiomerically pure, or be stereoisomeric or diastereomeric mixtures. In the case of amino acid residues, such residues can be of either the L- or D-form. The configuration for naturally occurring amino acid residues is generally L. When not specified the residue is the L form. It is to be understood that the chiral centers of the compounds provided herein can undergo epimerization *in vivo*. As such, one of skill in the art will recognize that administration of a compound in its (R) form is equivalent, for compounds that undergo epimerization *in vivo*, to administration of the compound in its (S) form.

The conjugates provided herein are prodrugs because they include a therapeutic agent in an inactive form that is ultimately converted to an active form at the targeted cell or tissue or in the environment thereof. Upon exposure to targeted protease either a biologically, pharmaceutically or therapeutically active form of a compound is released, or, a derivative that can be further metabolized into a biologically, pharmaceutically or therapeutically active form of a compound.

As used herein, substantially pure means sufficiently homogeneous to appear free of readily detectable impurities as determined by standard methods of analysis, such as thin layer chromatography (TLC), gel electrophoresis, high performance liquid chromatography (HPLC) and mass spectrometry (MS), used by those of skill in the art to assess such purity, or sufficiently pure such that further purification would not alter the physical and chemical properties, such as enzymatic and biological activities, of the substance for its intended purpose. Methods for purification of the compounds to produce substantially chemically pure compounds are known to those of skill in the art. A substantially



-44-

chemically pure compound may, however, be a mixture of stereoisomers. In such instances, further purification might increase the specific activity of the compound.

As used herein, a peptidic substrate includes peptides and molecules,  
 5 such as peptide mimetics and peptides that include peptide bond surrogates.

As used herein, conventional terminology (Schechter *et al.* (1967) *Biochem. Biophys. Res. Commun.* 27:157-162) is used to refer to specific subsites of a protease substrate:

$P_n \dots P_3-P_2-P_1 \downarrow P_1'-P_2'-P_3' \dots P_n'$ . The scissile bond (i.e., the cleavage site) of a  
 10 substrate is indicated by the arrow. Positions N-terminal of that bond are referred to as unprimed positions. Subsites are then assigned a number based on their distance from the scissile bond. Amino acids (or amino acid surrogates) that form the scissile bond are assigned the number 1, adjacent residues the number 2, and so on, counting away from the scissile bond. Each specific  
 15 subsite of the substrate, therefore, is uniquely identified by a number and the designation as primed or unprimed.

As used herein, a surrogate of a peptide bond is a divalent group that possesses similar steric and/or electronic characteristics to  $-C(O)NH-$ . Peptide bond surrogates include, but are not limited to, alkene isosteres ( $-CR=CR-$ ),  
 20 particularly (E)-alkene isosteres of formula  $-CH=CH-$ , hydroxyethylene isosteres ( $-CH(OH)CH_2-$ ), enamine isosteres ( $-C(=CRR)NH-$ ), aminoalcohol isosteres ( $-CH(OH)CH_2NH-$ ), difluoroketone isosteres ( $-C(O)CF_2-$ ), retroinverso compounds ( $-NHC(O)-$ ), divalent heterocyclyl or heteroaryl groups, and cyclopropyl isosteres  
 25 such as:

25



30

As used herein, alkyl, alkenyl and alkynyl carbon chains, if not specified, contain from 1 to 20 carbons, generally 1 to 16 carbons, and are straight or branched. Alkenyl carbon chains of from 2 to 20 carbons typically contain 1 to

-45-

8 double bonds, and the alkenyl carbon chains of 2 to 16 carbons and typically contain 1 to 5 double bonds. Alkynyl carbon chains of from 2 to 20 carbons typically contain 1 to 8 triple bonds, and the alkynyl carbon chains of 2 to 16 carbons and generally contain 1 to 5 triple bonds. Exemplary alkyl, alkenyl and alkynyl groups herein include, but are not limited to, methyl, ethyl, propyl, isopropyl, isobutyl, n-butyl, sec-butyl, tert-butyl, isopentyl, neopentyl, tert-pentyl and isohexyl. The alkyl, alkenyl and alkynyl groups, unless otherwise specified, optionally can be substituted, with one or more groups, generally alkyl group substituents that are the same or different. As used herein, lower alkyl, lower alkenyl, and lower alkynyl refer to carbon chains having less than about 6 carbons. As used herein, "alk(en)(yn)yl" refers to an alkyl group containing at least one double bond and at least one triple bond.

As used herein, "cycloalkyl" refers to a saturated mono- or multicyclic ring system, typically 3 to 10 carbon atoms, such as, for example, 3 to 6 carbon atoms; cycloalkenyl and cycloalkynyl refer to mono- or multicyclic ring systems that respectively include at least one double bond and at least one triple bond. Cycloalkenyl and cycloalkynyl groups contain, for example, 3 to 10 carbon atoms, with cycloalkenyl groups generally containing 4 to 7 carbon atoms and cycloalkynyl groups that contain, for example 8 to 10 carbon atoms. The ring systems of the cycloalkyl, cycloalkenyl and cycloalkynyl groups can be composed of one ring or two or more rings which can be joined together in a fused, bridged or spiro-connected fashion, and optionally can be substituted with one or more alkyl group substituents. "Cycloalk(en)(yn)yl" refers to a cycloalkyl group containing at least one double bond and at least one triple bond.

As used herein, "substituted alkyl," "substituted alkenyl," "substituted alkynyl," "substituted cycloalkyl," "substituted cycloalkenyl," and "substituted cycloalkynyl" refer to alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkenyl and cycloalkynyl groups, respectively, that are substituted with one or more substituents, in certain embodiments one to three substituents, independently selected from alkyl, halo, haloalkyl, such as halo lower alkyl, pseudohalo, aryl, amino, dialkylamino, nitro, cyano, azido, alkylsulfinyl, alkylsulfonyl,

-46-

alkylcarbonylamino, alkoxycarbonylamino, aminoimino, hydroxy, alkoxy, aryloxy, alkyloxy, alkylthio, arylthio, aralkyloxy, aralkylthio, carboxy, alkylcarbonyl, alkoxycarbonyl, oxo and cycloalkyl.

As used herein, "aryl" refers to cyclic groups containing from 6 to 19 carbon atoms. Aryl groups include, but are not limited to groups, such as fluorenyl, substituted fluorenyl, phenyl, substituted phenyl, naphthyl and substituted naphthyl. As used herein, "aryl" also refers to aryl-containing groups, including, but not limited to, aryloxy, arylthio, arylcarbonyl and arylamino groups.

As used herein, "heteroaryl" refers to a monocyclic or multicyclic aromatic ring system, generally about 5 to about 15 members where one or more, such as 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and sulfur atoms. The heteroaryl group optionally can be fused to a benzene ring. Exemplary heteroaryl groups include, for example, furyl, imidazolyl, pyrrolidinyl, pyrimidinyl, tetrazolyl, thienyl, pyridyl, pyrrolyl, N-methylpyrrolyl, quinolinyl and isoquinolinyl, with pyridyl, thienyl and quinolinyl as examples thereof.

As used herein, "heteroaryl" also refers to heteroaryl-containing groups, including, but not limited to, heteroaryloxy, heteroarylthio, heteroarylcarbonyl and heteroarylamino.

As used herein, "heterocyclyl" refers to a monocyclic or multicyclic non-aromatic ring system, such as systems of 3 to 10 members, for example 4 to 7 members or 5 to 6 members, where one or more, such as 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and/or sulfur atoms.

As used herein, "substituted aryl," "substituted heteroaryl" and "substituted heterocyclyl" refer to aryl, heteroaryl and heterocyclyl groups, respectively, that are substituted with one or more substituents, in certain embodiments one to three substituents, independently selected from alkyl, cycloalkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, such as 1 to 3, substituents selected from halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2

-47-

triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy, haloalkyl and polyhaloalkyl, such as halo lower alkyl, especially trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that optionally is substituted with 1 or more, generally 1 to 3, substituents selected from halo, halo alkyl and alkyl,

- 5 heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxy carbonyl, aminoimino, alkoxycarbonylamino, aryloxy carbonylamino, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, arylaminocarbonyl, diarylaminocarbonyl, aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkylaminoalkyl, dialkylaminoalkyl, arylaminoalkyl, amino, alkylamino, 10 dialkylamino, arylamino, alkylaryl amino, alkylcarbonylamino, arylcarbonylamino, azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyno, isothiocyno, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl, alkylaminosulfonyl, dialkylaminosulfonyl and arylaminosulfonyl.

- As used herein, "aralkyl" refers to an alkyl group in which one of the 15 hydrogen atoms of the alkyl is replaced by an aryl group.

As used herein, "heteroaralkyl" refers to an alkyl group in which one of the hydrogen atoms of the alkyl is replaced by a heteroaryl group.

- As used herein, the nomenclature alkyl, alkoxy, carbonyl, etc. is used as is generally understood by those of skill in this art. For example, as used herein 20 alkyl refers to saturated carbon chains that contain one or more carbons; the chains can be straight or branched or include cyclic portions or be cyclic.

- Where the number of any given substituent is not specified (*e.g.*, "haloalkyl"), there can be one or more substituents present. For example, "haloalkyl" can include one or more of the same or different halogens. As 25 another example, "C<sub>1-3</sub>alkoxyphenyl" can include one or more of the same or different alkoxy groups containing one, two or three carbons.

As used herein, "halo", "halogen" or "halide" refers to F, Cl, Br or I.

- As used herein, pseudohalides are compounds that behave substantially similar to halides. Such compounds can be used in the same manner and 30 treated in the same manner as halides (X<sup>-</sup>, in which X is a halogen, such as Cl or Br). Pseudohalides include, but are not limited to, cyanide, cyanate,

-48-

thiocyanate, selenocyanate, trifluoromethoxy, difluoromethoxy, dichloromethoxy and azide.

As used herein, "haloalkyl" refers to a lower alkyl radical in which one or more of the hydrogen atoms are replaced by halogen. Such groups include, but not limited to, chloromethyl, trifluoromethyl, 1-chloro-2-fluoroethyl and the like.

As used herein, "haloalkoxy" refers to RO- in which R is a haloalkyl group.

As used herein, "sulfinyl" or "thionyl" refers to -S(O)-. As used herein, "sulfonyl" or "sulfuryl" refers to -S(O)<sub>2</sub>-. As used herein, "sulfo" refers to -S(O)<sub>2</sub>O-.

As used herein, "carboxy" refers to a divalent radical, -C(O)O-.

As used herein, "aminocarbonyl" refers to -C(O)NH<sub>2</sub>.

As used herein, "alkylaminocarbonyl" refers to -C(O)NHR in which R is hydrogen or alkyl, such as, for example, lower alkyl.

As used herein "dialkylaminocarbonyl" as used herein refers to -C(O)NR'R in which R' and R are independently selected from hydrogen or alkyl, such as, for example, lower alkyl; "carboxamide" refers to groups of formula -NR'COR.

As used herein, "diarylaminocarbonyl" refers to -C(O)NRR' in which R and R' are independently selected from aryl, such as lower aryl, for example, phenyl.

As used herein, "aralkylaminocarbonyl" refers to -C(O)NRR' in which one of R and R' is aryl, such as, lower aryl, for example, phenyl, and the other of R and R' is alkyl, such as, for example, lower alkyl.

As used herein, "arylaminocarbonyl" refers to -C(O)NHR in which R is aryl, such as lower aryl, for example, phenyl.

As used herein, "hydroxycarbonyl" refers to -COOH.

As used herein, "alkoxycarbonyl" refers to -C(O)OR in which R is alkyl, such as lower alkyl.

As used herein, "aryloxycarbonyl" refers to -C(O)OR in which R is aryl, such lower aryl, for example phenyl.

As used herein, "alkoxy" and "alkylthio" refer to RO- and RS-, in which R is alkyl, such as, for example, lower alkyl.

As used herein, "aryloxy" and "arylthio" refer to RO- and RS-, in which R is aryl, such lower aryl, for example, phenyl.

As used herein, "alkylene" refers to a straight, branched or cyclic, such as, for example, straight or branched, divalent aliphatic hydrocarbon group, for example, having from 1 to about 20 carbon atoms such as 1 to 12 carbons, and for example, is lower alkylene. There optionally can be inserted along the alkylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkylene groups include methylene (-CH<sub>2</sub>-), ethylene (-CH<sub>2</sub>CH<sub>2</sub>-), propylene (-(CH<sub>2</sub>)<sub>3</sub>-), cyclohexylene (-C<sub>6</sub>H<sub>10</sub>-), methylenedioxy (-O-CH<sub>2</sub>-O-) and ethylenedioxy (-O-(CH<sub>2</sub>)<sub>2</sub>-O-). The term "lower alkylene" refers to alkylene groups having 1 to 6 carbons. Exemplary alkylene groups are lower alkylene, such as, for example, alkylene of 1 to 3 carbon atoms.

As used herein, "alkenylene" refers to a straight, branched or cyclic, typically straight or branched, divalent aliphatic hydrocarbon group, such as, for example, having from 2 to about 20 carbon atoms and at least one double bond, generally 1 to 12 carbons, and is for example, lower alkenylene. There optionally can be inserted along the alkenylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkenylene groups include -CH=CH-CH=CH- and -CH=CH-CH<sub>2</sub>-. The term "lower alkenylene" refers to alkenylene groups having 2 to 6 carbons. Exemplary alkenylene groups are lower alkenylene, such as, for example, alkenylene of 3 to 4 carbon atoms.

As used herein, "alkynylene" refers to a straight, branched or cyclic, generally straight or branched, divalent aliphatic hydrocarbon group, such those having from 2 to about 20 carbon atoms and at least one triple bond, generally 1 to 12 carbons, such as, for example, lower alkynylene. There optionally can be inserted along the alkynylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkynylene groups include -C≡C-C≡C-, -C≡C- and -C≡C-CH<sub>2</sub>-. The term "lower alkynylene" refers to

-50-

alkynylene groups having 2 to 6 carbons. Exemplary alkynylene groups are lower alkynylene, such as, for example, alkynylene of 3 to 4 carbon atoms.

As used herein, "alk(en)(yn)ylene" refers to a straight, branched or cyclic, generally straight or branched, divalent aliphatic hydrocarbon group, having, for example, from 2 to about 20 carbon atoms and at least one triple bond, and at least one double bond; typically 1 to 12 carbons, such as, for example, lower alk(en)(yn)ylene. There optionally can be inserted along the alkynylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alk(en)(yn)ylene groups include  $\text{—C=C—(CH}_2\text{)}_n\text{—C}\equiv\text{C—}$ , where n is 1 or 2. The term "lower alk(en)(yn)ylene" refers to alk(en)(yn)ylene groups having up to 6 carbons. Exemplary alk(en)(yn)ylene groups are lower alk(en)(yn)ylene, such as, for example, alk(en)(yn)ylene of 4 carbon atoms.

As used herein, "cycloalkylene" refers to a divalent saturated mono- or multicyclic ring system, generally 3 to 10 carbon atoms, such as 3 to 6 carbon atoms; cycloalkenylene and cycloalkynylene refer to divalent mono- or multicyclic ring systems that respectively include at least one double bond and at least one triple bond. Cycloalkenylene and cycloalkynylene groups can contain 3 to 10 carbon atoms, with, for example, cycloalkenylene groups containing 4 to 7 carbon atoms and cycloalkynylene groups containing 8 to 10 carbon atoms. The ring systems of the cycloalkylene, cycloalkenylene and cycloalkynylene groups can be composed of one ring or two or more rings that can be joined together in a fused, bridged or spiro-connected fashion. "Cycloalk(en)(yn)ylene" refers to a cycloalkylene group containing at least one double bond and at least one triple bond.

As used herein, "substituted alkylene," "substituted alkenylene," "substituted alkynylene," "substituted cycloalkylene," "substituted cycloalkenylene," and "substituted cycloalkynylene" refer to alkylene, alkenylene, alkynylene, cycloalkylene, cycloalkenylene and cycloalkynylene groups, respectively, that are substituted with one or more substituents, in certain embodiments one to three substituents, independently selected from halo, haloalkyl, such as, for example, halo lower alkyl, aryl, hydroxy, alkoxy, aryloxy,

alkyloxy, alkylthio, arylthio, aralkyloxy, aralkylthio, carboxy alkoxy carbonyl, oxo and cycloalkyl.

As used herein, "arylene" refers to a monocyclic or polycyclic, such as monocyclic, divalent aromatic group, for example, having from 5 to about 20  
5 carbon atoms and at least one aromatic ring, such as 5 to 12 carbons, and, is, for example, lower arylene. There optionally can be inserted around the arylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary arylene groups include 1,2-, 1,3- and 1,4-phenylene. The term  
10 "lower arylene" refers to arylene groups having 5 or 6 carbons. Exemplary arylene groups are lower arylene.

As used herein, "heteroarylene" refers to a divalent monocyclic or multicyclic aromatic ring system, such as of about 5 to about 15 members where one or more, typically, for example, 1 to 3 of the atoms in the ring  
15 system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and/or sulfur atom(s).

As used herein, "heterocyclylene" refers to a divalent monocyclic or multicyclic non-aromatic ring system, generally of 3 to 10 members, such as, for example, 4 to 7 members or 5 to 6 members, where one or more, such as, for  
20 example, 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and/or sulfur atom(s).

As used herein, "substituted arylene," "substituted heteroarylene" and "substituted heterocyclylene" refer to arylene, heteroarylene and heterocyclylene groups, respectively, that are substituted with one or more substituents, in  
25 certain embodiments one to three substituents, independently selected from alkyl, cycloalkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, such as 1 to 3, substituents selected from halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2 triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy,  
30 haloalkyl and polyhaloalkyl, such as, halo lower alkyl, for example trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that optionally is substituted with 1 or more, such as 1 to 3, substituents selected from, for example, halo,



halo alkyl and alkyl, heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxy, aryloxy, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, arylaminocarbonyl, diarylaminocarbonyl, aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkylaminoalkyl, dialkylaminoalkyl, arylaminoalkyl, amino, alkylamino, dialkylamino, arylamino, alkylaryl amino, alkylcarbonylamino, arylcarbonylamino, azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyno, isothiocyno, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl, alkylaminosulfonyl, dialkylaminosulfonyl and arylaminosulfonyl.

- 10 As used herein, "alkylidene" refers to a divalent group, such as  $=CR'R''$ , which is attached to one atom of another group, forming a double bond. Exemplary alkylidene groups are methylenes ( $=CH_2$ ) and ethylenes ( $=CHCH_3$ ). As used herein, "aralkylidene" refers to an alkylidene group in which either  $R'$  or  $R''$  is an aryl group. "Cycloalkylidene" groups are those where  $R'$  and  $R''$  are
- 15 linked to form a carbocyclic ring. "Heterocyclidene" groups are those where at least one of  $R'$  and  $R''$  contain a heteroatom in the chain, and  $R'$  and  $R''$  are linked to form a heterocyclic ring.

- As used herein, "amido" refers to the divalent group  $-C(O)NH-$ . "Thioamido" refers to the divalent group  $-C(S)NH-$ . "Oxyamido" refers to the
- 20 divalent group  $-OC(O)NH-$ . "Thiaamido" refers to the divalent group  $-SC(O)NH-$ . "Dithiaamido" refers to the divalent group  $-SC(S)NH-$ . "Ureido" refers to the divalent group  $-HNC(O)NH-$ . "Thioureido" refers to the divalent group  $-HNC(S)NH-$ .

- As used herein, "semicarbazide" refers to  $-NHC(O)NHNH-$ . "Carbazate" refers to the divalent group  $-OC(O)NHNH-$ . "Isothiocarbazate" refers to the
- 25 divalent group  $-SC(O)NHNH-$ . "Thiocarbazate" refers to the divalent group  $-OC(S)NHNH-$ . "Sulfonylhydrazide" refers to the group  $-SO_2NHNH-$ . "Hydrazide" refers to the divalent group  $-C(O)NHNH-$ . "Azo" refers to the divalent group  $-N=N-$ . "Hydrazinyl" refers to the divalent group  $-NH-NH-$ .

- 30 As used herein, the term "amino acid" refers to  $\alpha$ -amino acids which are racemic, or of either the D- or L-configuration. The designation "d" preceding an amino acid designation (e.g., dAla, dSer, dVal, etc.) refers to the D-isomer of the

-53-

amino acid. The designation "dl" preceding an amino acid designation (*e.g.*, dlPip) refers to a mixture of the L- and D-isomers of the amino acid.

As used herein, when any particular group, such as phenyl or pyridyl, is specified, this means that the group is unsubstituted or is substituted.

- 5 Exemplary substituents where not specified are halo, halo lower alkyl, and lower alkyl.

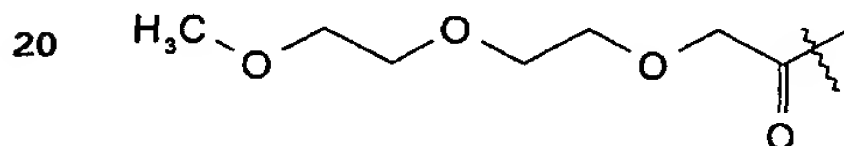
As used herein, the abbreviations for any protective groups, amino acids and other compounds, are, unless indicated otherwise, in accord with their common usage, recognized abbreviations, or the IUPAC-IUB Commission on

- 10 Biochemical Nomenclature (see, (1972) *Biochem.* 11:942-944).

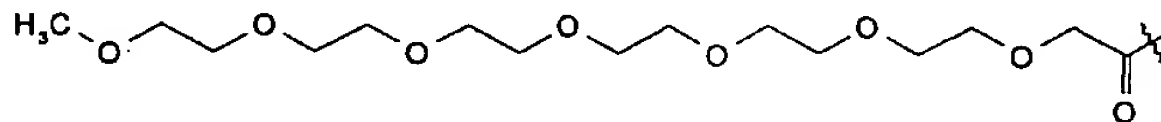
As used herein, HHT and CHT refer to hexahydrotyrosyl (also known as cyclohexyltyrosyl or p-hydroxycyclohexylalanyl), CHA is cyclohexylalanyl, Pyr and pyroGlu refer to pyroglutamic acid, Pip is pipercolinic acid, Sar is sarcosine, nLeu and Nle are norleucine, nVal is norvaline, Aib is 2-aminoisobutyric acid,

- 15 Quat is (R)-Glu( $\alpha$ -(3-amidinobenzyl)), and Abu and But are 2-aminobutyric acid.

As used herein, PEG represents a polyethylene glycol containing substituent having the designated number of ethyleneoxy subunits. Thus, the term PEG(2) represents:



- 25 and the term PEG(6) represents:

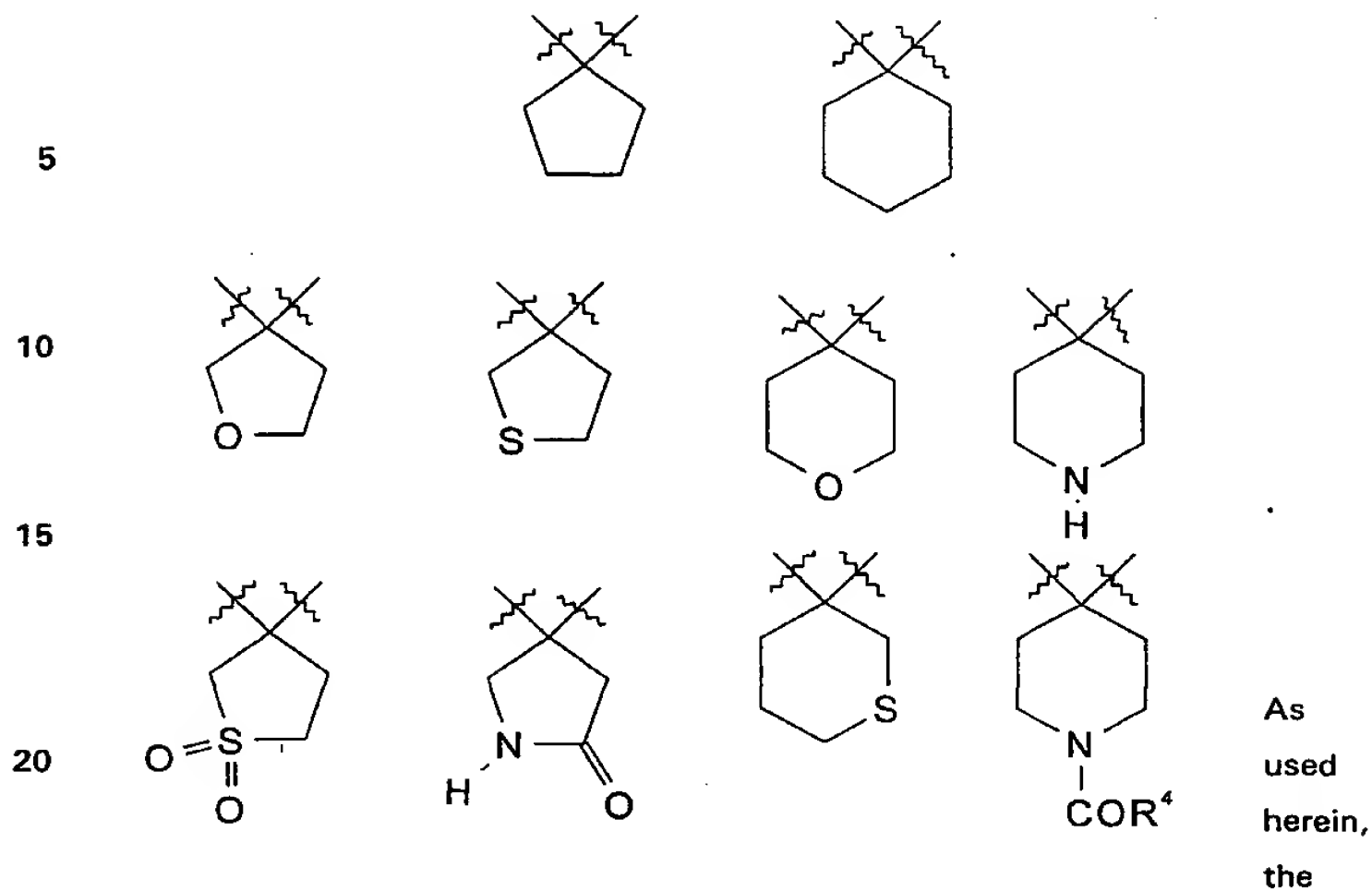


30

When R<sup>1</sup> and R<sup>2</sup> are combined to form -(CH<sub>2</sub>)<sub>n</sub>-, the cyclic moieties and heteroatom-containing cyclic moieties so defined include, but are not limited to:

35

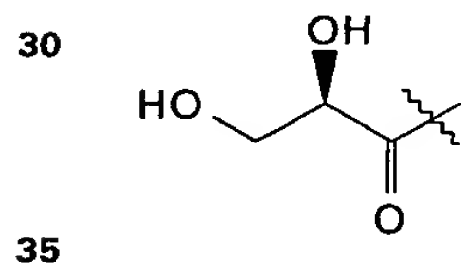
-54-



term "hydroxylated" represents substitution on a substitutable carbon of the ring system being so described by a hydroxyl moiety. As used herein, the term "poly-

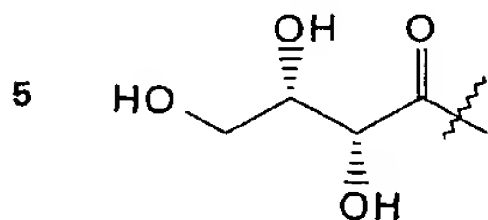
25 hydroxylated" represents substitution on two or more substitutable carbons of the ring system being so described by 2, 3 or 4 hydroxyl moieties.

As used herein, the term "(d)(2,3-dihydroxypropionyl)" represents the following structure:

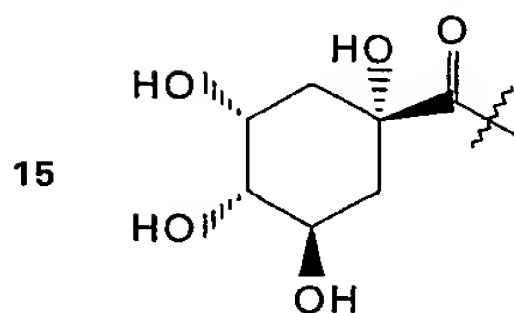


-55-

As used herein, the term "(2R,3S)-2,3,4-trihydroxybutanoyl" represents the following structure:

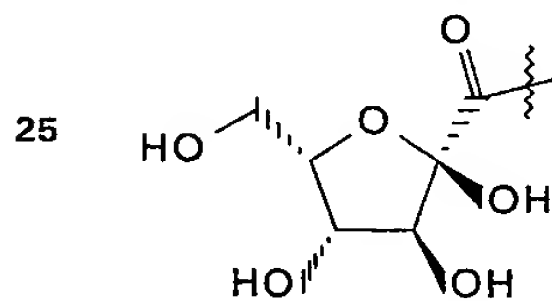


10 As used herein, the term "quinylyl" represents the following structure:



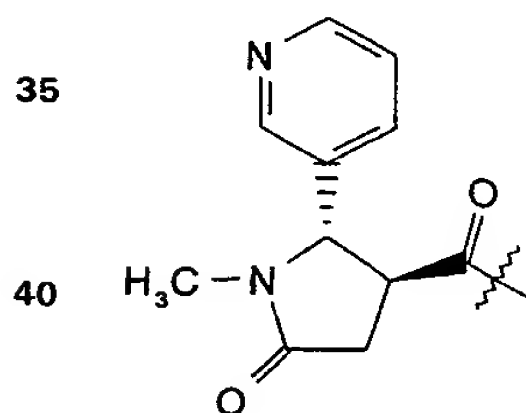
20 or a diastereomer thereof.

As used herein, the term "gulonyl" represents the following structure:



30 or a diastereomer thereof.

As used herein, the term "cotininylyl" represents the following structure:



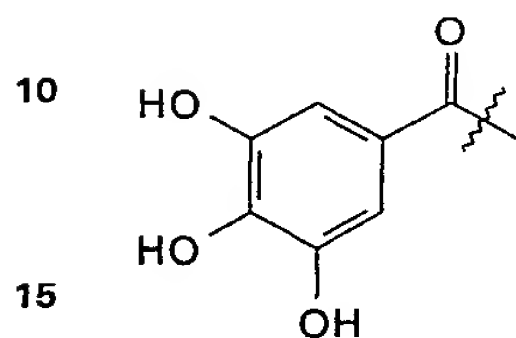
45

-56-

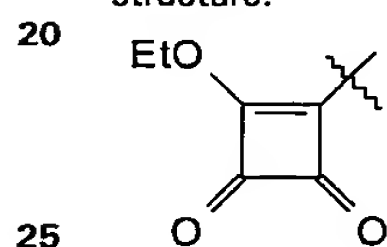
5

or a diastereomer thereof.

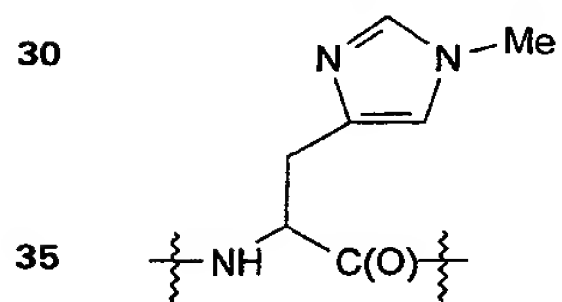
As used herein, the term "gallyl" represents the following structure:



As used herein, the term "4-ethoxysquaryl" represents the following structure:



As used herein, 1-methylHis or (1Me)H refers to the structure:

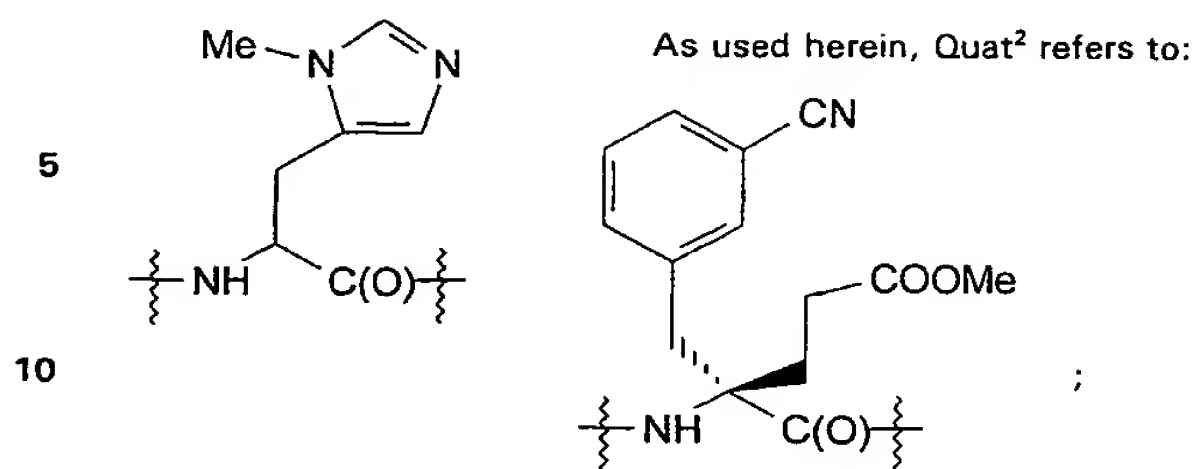


As used herein, 3-methylHis or (3Me)H refers to the structure:

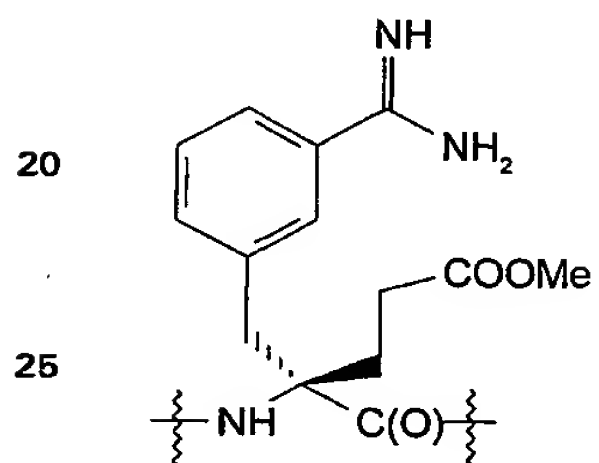
40

45

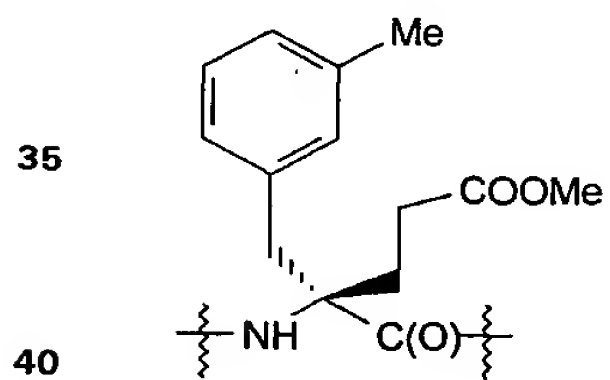
-57-



15 Quat<sup>3</sup> refers to:



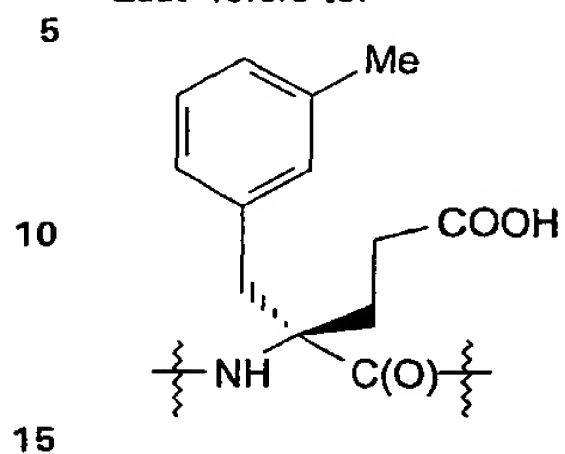
30 Quat<sup>4</sup> refers to:



45

-58-

; and

Quat<sup>5</sup> refers to:

Other abbreviations as used herein are as follows:

Abbreviation	Refers to
Aib	2-aminoisobutyryl
20 4,4-dimethylThr	2-amino-3-hydroxy-4-methylpentanoyl
Met(O <sub>2</sub> )	methioninyl-S,S-dioxide
Ser(OMe)	the O-methyl ether of serinyl, also known as 2-amino-3-methoxypropanoyl
hSer	homoserinyl, also known as 2-amino-4-hydroxybutanoyl
25 (hS)Gly	N-(2-hydroxyethyl)glycyl
N,N-dimethylGly	N,N-dimethylglycyl
β-Ala	3-aminopropanoyl
Cys(Me)	S-methylcysteinyl
30 t-butylGly	2-amino-3,3-dimethylbutanoyl
F(Gn)	4-guanidinylphenylalanyl
hCHA	homocyclohexylalanyl, or 2-amino-4-cyclohexylbutanoyl
hexylGly	2-amino-octanoyl
35 allylGly	2-amino-4-pentenoyl
Inact.	inactive
NT	not tested

-59-

	MeOEtCO	3-methoxypropanoyl
	3,4-MethyldioxyPhAc	3,4-methylenedioxyphenylacetyl
	L-3-PhLactyl	L-2-hydroxy-3-phenylpropanoyl
	MeOEtOCO	2-methoxyethoxycarbonyl
5	MeOCO	methoxycarbonyl
	MeO(EtO)2Ac	2-(2-methoxyethoxy)ethoxyacetyl
	2-PyridylAc	2-pyridylacetyl
	PhOAc	phenoxyacetyl
	MeOAc	methoxyacetyl
10	PhAc	phenylacetyl
	MeOEtOAc	2-methoxyethoxyacetyl
	HOOCButa	glutaryl
	Z	benzyloxycarbonyl
	EtOCO	ethoxycarbonyl
15	$\beta$ A	beta-alanyl or 3-aminopropanoyl
	NapAc	1-naphthylacetyl
	iBoc	isobutoxycarbonyl
	HOAc	hydroxyacetyl
	MeSucc	3-methoxycarbonylpropanoyl
20	Succ	succinyl
	HCO	formyl
	4-(guan)Phg	4-guanidinyphenylglycyl
	Dox	doxorubicin
	Tax	taxol
25	dA(Chx) or dCha	d-cyclohexylalanyl
	dhF	d-homophenylalanyl
	P(OH)	4-hydroxypropyl

**B. Protease targets**

30 The conjugates herein are designed to target proteases that are located on cell surfaces, particularly tumor cells and cells involved in tumorigenic processes and angiogenesis and other proliferative processes. The conjugates, described in detail below, contain a peptidic substrate for a selected targeted



-60-

cell surface protease linked, either directly or via a linker, to a therapeutic agent, typically a cytotoxic agent, which is substantially inactive when in the conjugate. The therapeutic agent is released in a form that is active or that can be activated in the vicinity of the targeted cell or tissue to which it is delivered.

- 5 As a result, active therapeutic agent accumulates at the targeted cells or tissue or in the targeted cells.

The targeted protease is selected by identifying a protease that is located on a cell or tissue (or associated therewith) that is involved in the disease process or serendipitously present in the locale of cells or tissues involved in the disease or disease process, and, generally, is not located at all or present or  
10 active at lower levels, generally substantially lower levels, or exhibits altered activity or specificity, on many, if not all, other cells or tissues. The variety and numbers of non-targeted cells or tissues that express the active protease varies for particular proteases and diseases intended for treatment. Those of skill in  
15 the art will select a target based upon the disease, targeted agents and tolerable or acceptable levels of side-effects. The goal is to achieve enhanced therapeutic index compared with administration of the targeted agent by itself.

The targeted protease may or may not be involved in the disease process and its expression can be serendipitous; for purposes herein its particular role or  
20 lack thereof is not important; it is the fact that it is active in the locale of targeted tissues or cells that is important. For example, many of the cell surface proteases of interest herein are expressed or active on tumor cells or cells involved in the tumorigenic processes. Any method known to one of skill in the art for determining or detecting a tissue or cell expression profile can be used.  
25 For example, RNA blots composed of RNA from numerous tissues (*e.g.*, a multiple tissue expression (MTE) array available from CLONTECH, Palo Alto, CA), can be screened with probes based upon the nucleic acid sequence of the protease of interest to identify cells that express the protease. Northern analysis of the blots to test for expression also can be used.

30 Included among the targeted proteases are those designated type II membrane-bound serine proteases (MTSPs; see, *e.g.*, U.S. application Serial No. 09/776,191, filed February 2, 2001 and International PCT application No.

-61-

PCT/US01/03471 published as International PCT application No. WO 01/57194; see International PCT application No. PCT/US02/07903; see, also U.S. provisional application Serial Nos. 60/275,592, 60/278,166, 60/279,228, 60/291,001, 60/291,501 60/316,818, 60/302,939, 60/316,818, 60/328,529, 5 60/328,530, 60/332,015, 60/328,939, and provisional application, filed on May 20, 2002 under attorney docket no. 24745-P1624; U.S. application Serial Nos. 10/099,700, 10/104,271, 10/112,221, application filed on May 14, 2002 under attorney docket no. 24745-1616) and those found on endothelial cells designated endotheliases (see, U.S. application Serial No. 09/717,473, filed 10 November 20, 2000, and International PCT application No. PCT/US00/31803 published as International PCT application No. WO 01/36604); see, also SEQ ID Nos. 3-26, 269-270 and 272-276.

Also contemplated are proteases that are located at the cell surface by virtue of a specific interaction with a cell surface protein. Urokinase 15 plasminogen activator (u-PA) bound to urokinase plasminogen activator receptor (u-PAR) is exemplary of such proteases. Nucleic acid sequence information and expression profiles of exemplary MTSPs and endotheliases are as follows (see, also EXAMPLE 6).

#### 1. MTSPs

20 Cell surface proteolysis is a mechanism for the generation of biologically active proteins that mediate a variety of cellular functions. These membrane-anchored proteins, include a disintegrin-like and metalloproteinase (ADAM) and membrane-type matrix metalloproteinase (MT-MMP). In addition to the MMPs, serine proteases have been implicated in neoplastic disease progression. Most 25 serine proteases, which are either secreted enzymes or are sequestered in cytoplasmic storage organelles, have roles in blood coagulation, wound healing, digestion, immune responses and tumor invasion and metastasis.

Transmembrane serine proteases (MTSPs) appear to be involved in the etiology and pathogenesis of tumors. These enzymes are expressed in certain 30 cancerous and tumor cells and in other cells associated with other proliferative disorders and other disease states, such as in inflammatory cells and can be tissue or organ-specific. In mammals, more than 20 members of the family are

-62-

known (see, Hooper *et al.* (2001) *J. Biol. Chem.* 276:857-860, see, also U.S. application Serial No. 09/776,191, filed February 2, 2001 and International PCT application No. PCT/US01/03471; see, also U.S. provisional application Serial Nos. 60/275,592 and 60/278,166; and see SEQ ID Nos. 1-37). These include

5 corin (accession nos. AF133845 and AB013874; see, Yan *et al.* (1999) *J. Biol. Chem.* 274:14926-14938; Tomia *et al.* (1998) *J. Biochem.* 124:784-789; Uan *et al.* (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:8525-8529); enterpeptidase (also designated enterokinase; accession no. U09860 for the human protein; see, Kitamoto *et al.* (1995) *Biochem.* 27: 4562-4568; Yahagi *et al.* (1996) *Biochem.*

10 *Biophys. Res. Commun.* 219:806-812; Kitamoto *et al.* (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91:7588-7592; Matsushima *et al.* (1994) *J. Biol. Chem.* 269:19976-19982;); human airway trypsin-like protease (HAT; accession no. AB002134; see Yamaoka *et al.* *J. Biol. Chem.* 273:11894-11901); MTSP1 (also called TADG-15 and matriptase, see SEQ ID Nos. 1 and 2; accession nos.

15 AF133086/AF118224, AF04280022; Takeuchi *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:11054-1161; Lin *et al.* (1999) *J. Biol. Chem.* 274:18231-18236; Takeuchi *et al.* (2000) *J. Biol. Chem.* 275:26333-26342; and Kim *et al.* (1999) *Immunogenetics* 49:420-429); hepsin (see, accession nos. M18930, AF030065, X70900; Leytus *et al.* (1988) *Biochem.* 27: 11895-11901; Vu *et al.* (1997) *J. Biol. Chem.* 272:31315-31320; and Farley *et al.* (1993) *Biochem.*

20 *Biophys. Acta* 1173:350-352; and see, U.S. Patent No. 5,972,616); TMPRS2 (see, Accession Nos. U75329 and AF113596; Paoloni-Giacobino *et al.* (1997) *Genomics* 44:309-320; and Jacquinet *et al.* (2000) *FEBS Lett.* 468: 93-100); and TMPRSS4 (see, Accession No. NM 016425; Wallrapp *et al.* (2000) *Cancer*

25 60:2602-2606). Also known MTSP3, MTSP4, MTSP6, MTSP7, MTSP9, MTSP10, MTSP12, MTSP20, MTSP22 and MTSP25 (see, SEQ ID NOs. 3-26, 269-270 and 272-276; see, also U.S. application Serial No. 09/776,191, filed February 2, 2001 and International PCT application No. PCT/US01/03471 published as International PCT application No. WO 01/57194; see International

30 PCT application No. PCT/US02/07903; see, also U.S. provisional application Serial Nos. 60/275,592, 60/278,166, 60/279,228, 60/291,001, 60/291,501 60/316,818, 60/302,939, 60/316,818, 60/328,529, 60/328,530,

-63-

60/332,015, 60/328,939, and provisional application, filed on May 20 2002, under attorney docket no. 24745-P1624; U.S. application Serial Nos. 10/099,700, 10/104,271, 10/112,221, application filed on May 14, 2002 under attorney docket no. 24745-1616)).

5           Serine proteases, including transmembrane serine proteases, have been implicated in processes involved in neoplastic development and progression. While the precise role of these proteases has not been elaborated, serine proteases and inhibitors thereof are involved in the control of many intra- and extracellular physiological processes, including degradative actions in cancer cell  
10 invasion, metastatic spread, and neovascularization of tumors, that are involved in tumor progression. It is believed that proteases are involved in the degradation of extracellular matrix (ECM) and contribute to tissue remodeling, and are necessary for cancer invasion and metastasis. The activity and/or expression of some proteases have been shown to correlate with tumor  
15 progression and development, and also are shown to be active in specific cell types.

For example, a membrane-type serine protease MTSP1 (also called matriptase; see SEQ ID Nos. 1 and 2 from U.S. Patent No. 5,972,616; and GenBank Accession No. AF118224; (1999) *J. Biol. Chem.* 274:18231-18236;  
20 U.S. Patent No. 5,792,616; see, also Takeuchi (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:11054-1161) that is expressed in epithelial cancer and normal tissue (Takeuchi *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:11054-61) has been identified. It has been proposed that it plays a role in the metastasis of breast cancer. Its primary cleavage specificity is Arg-Lys residues. Matriptase also is  
25 expressed in a variety of epithelial tissues with high levels of activity and/or expression in the human gastrointestinal tract and the prostate.

Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer (Tanimoto *et al.* (1997) *Cancer Res.*, 57:2884-7). The hepsin transcript appears to be abundant in carcinoma tissue  
30 and is almost never expressed in normal adult tissue, including normal ovary. It has been suggested that hepsin is frequently overexpressed in ovarian tumors

-64-

and therefore can be a candidate protease in the invasive process and growth capacity of ovarian tumor cells.

A serine protease-like gene, designated normal epithelial cell-specific 1 (NES1) (Liu *et al.* (1996) *Cancer Res.* 56:3371-9) has been identified. Although  
5 expression of the NES1 mRNA is observed in all normal and immortalized nontumorigenic epithelial cell lines, the majority of human breast cancer cell lines show a drastic reduction or a complete lack of its expression. The structural similarity of NES1 to polypeptides known to regulate growth factor activity and a negative correlation of NES1 expression with breast oncogenesis suggest a  
10 direct or indirect role for this protease-like gene product in the suppression of tumorigenesis.

#### Exemplary MTSPs

Each MTSP has a characteristic tissue expression profile; the MTSPs in particular, although not exclusively expressed or activated in tumors, exhibit  
15 characteristic tumor tissue expression or activation profiles. In some instances, MTSPs can have different activity in a tumor cell from a non-tumor cell by virtue of a change in a substrate or cofactor therefor or other factor that would alter functional activity of the MTSP. Hence each can serve as a diagnostic marker for particular tumors, by virtue of a level of activity and/or expression or  
20 function in a subject (i.e. a mammal, particularly a human) with neoplastic disease, compared to a subject or subjects that do not have the neoplastic disease. In addition, detection of activity (and/or expression) in a particular tissue can be indicative of neoplastic disease. Also, by virtue of the activity and/or expression profiles of each, they can serve as therapeutic targets, such  
25 as by administration of modulators of the activity thereof, or, as by administration of a prodrug specifically activated by one of the MTSPs. Each or any of the MTSPs can exhibit activity or expression levels or substrate specificities that differ in tumor cells from the levels in normal cells. Such tumor cells include, but are not limited to, colon, lung, prostate, breast, esophagous,  
30 pancreas, cervix, uterus, endometrium, and other solid tumors and in blood and lymphatic tumors. Hence, conjugates provided herein can be designed by

-65-

selection of substrate specificity for treatment of any of such tumors and neoplastic conditions.

#### **Tissue expression profiles**

The following are exemplary tissue and gene (see also, EXAMPLE 8)

- 5 profiles of some exemplary MTSPs. These profiles are not intended to define the full scope of expression or activation of these MTSPs, but demonstrate that MTSPs are expressed in tumors, and, hence their expression or activation or substrate specificity on the surface of tumor cells can be exploited in the methods herein and conjugates, designed in accord with the methods herein and  
10 as exemplified herein, that are cleaved by one or more of these MTSPs can be prepared and employed for treatment of neoplastic or other diseases or conditions or to target to cells that express these proteins on their surfaces.

#### **MTSP1 (matriptase)**

- MTSP1 (also called matriptase) is a trypsin-like serine protease with  
15 broad spectrum cleavage activity and two potential regulatory modules. It was named "matriptase" based on its ability to degrade the extra-cellular matrix and its trypsin-like activity. When isolated from breast cancer cells (or T-47D cell conditioned medium), MTSP1 has been reported to be primarily in an uncomplexed form. MTSP1 has been isolated from human milk; when isolated  
20 from human milk, it was reported to be in one of two complexed forms, 95 kDa (the predominant form) and 110 kDa; uncomplexed MTSP1 was not detected (Liu, *et al.* (1999) *J. Biol. Chem.* 274:18237-18242). It has been proposed that MTSP1 exists as an uncomplexed protease when in its active state. In breast milk, it has been reported to exist in complex with a fragment of hepatocyte  
25 growth factor inhibitor-1 (HAI-1), a Kunitz-type serine protease inhibitor having activity against trypsin-like serine proteases.

- Nucleic acids encoding the protein designed matriptase were cloned from T-47D human breast cancer cell-conditioned medium (Lin *et al.* (1999) *J. Biol. Chem.* 274:18231-18236). Upon analysis of the cDNA, it was determined that  
30 the full length protease has 683 amino acids and contains three main structural regions: a serine protease domain near the carboxyl-terminal region, four tandem low-density lipoprotein receptor domains, and two tandem complement

-66-

subcomponents C1r and C1s (see SEQ ID No. 1). Studies to identify additional serine proteases made by cancer cells were done using PC-3 cells. A serine protease termed "MT-SP1" (MTSP1) by the authors, reported to be a transmembrane protease was cloned (Takeuchi *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:11054-11061). It was subsequently found that originally identified matriptase sequence is included in the translated sequence of the cDNA that encodes MTSP1. The nucleic acid encoding the protein originally designated matriptase is a partial MTSP1 clone that lacks 516 of the coding nucleotides (Takeuchi, *et al.*, *J. Biol. Chem* 275:26333-26342 (2000).) Since the reported matriptase encoding cDNA sequence encoded a possible initiating methionine, it was proposed that alternative splicing could yield a protein lacking the N-terminal region of MTSP1. Hence, matriptase herein is a variant form of MTSP1.

MTSP1 demonstrates trypsin-like protease activity and is a Type II transmembrane protein with an extracellular protease domain. Studies of substrate specificity of MTSP1 reveal that protease-activated receptor 2 (PAR2), pro-hepatocyte growth factor (pro-HGF) and single-chain urokinase-type plasminogen activator (sc-uPA) are macromolecular substrates of MTSP1. PAR2 functions in inflammation, cytoprotection and/or cell adhesion, while sc-uPa functions in tumor cell invasion and metastasis. HGF serves a growth and pro-angiogenic factor.

An exemplary nucleotide sequence encoding a human MTSP1 is set forth in SEQ ID Nos 1 and 2. As previously noted SEQ ID No. 1 sets for an MTSP1-encoding nucleic acid sequence. This sequence is the longer version and includes the protease domain, which is common to both variants.

MTSP1 is expressed in breast, prostate and colorectal tumors. Hence conjugates with substrates therefor can be used for treatment of such tumors.

### MTSP3

The MTSP3 transcript was detected in lung carcinoma (LX-1), colon adenocarcinoma (CX-1), colon adenocarcinoma (GI-112) and ovarian carcinoma (GI-102). No apparent signal was detected in another form of lung carcinoma

-67-

(GI-117), breast carcinoma (GI-101), pancreatic adenocarcinoma (GI-103) and prostatic adenocarcinoma (PC3).

#### MTSP4

The MTSP4 transcript, a DNA fragment encoding part of the LDL receptor domain and the protease domain was used to probe an RNA blot composed of 76 different human tissues (catalog number 7775-1; human multiple tissue expression (MTE) array; CLONTECH). As in the northern analysis of gel blot, a very strong signal was observed in the liver. Signals in other tissues were observed in (decreasing signal level): fetal liver > heart = kidney = adrenal gland = testis = fetal heart and kidney = skeletal muscle = bladder = placenta > brain = spinal cord = colon = stomach = spleen = lymph node = bone marrow = trachea = uterus = pancreas = salivary gland = mammary gland = lung. MTSP4 also is expressed less abundantly in several tumor cell lines including HeLa S3 = leukemia K-562 = Burkitt's lymphomas (Raji and Daudi) = colorectal adenocarcinoma (SW480) > lung carcinoma (A549) = leukemia MOLT-4 = leukemia HL-60. PCR of the MTSP4 transcript from cDNA libraries made from several human primary tumors xenografted in nude mice (human tumor multiple tissue cDNA panel, catalog number K1522-1, CLONTECH) was performed using MTSP4-specific primers. The MTSP4 transcript was detected in breast carcinoma (GI-101), lung carcinoma (LX-1), colon adenocarcinoma (GI-112) and pancreatic adenocarcinoma (GI-103). No apparent signal was detected in another form of lung carcinoma (GI-117), colon adenocarcinoma (CX-1), ovarian carcinoma (GI-102). and prostatic adenocarcinoma (PC3). The MTSP4 transcript was also detected in LNCaP and PC-3 prostate cancer cell lines as well as in HT-1080 human fibrosarcoma cell line.

#### MTSP6

MTSP6 is expressed at high levels in the colon. It also is expressed in the, stomach, trachea, mammary gland, thyroid gland, salivary gland, pituitary gland and pancreas. It is expressed at lower levels in other tissues (see EXAMPLE 6).



-68-

MTSP6 also is expressed in several tumor cell lines including HeLa S3 > colorectal adenocarcinoma (SW480) > leukemia MOLT-4 > leukemia K-562. In mouse xenograft models, the MTSP6 transcript was strongly detected in lung carcinoma (LX-1), moderately detected in pancreatic adenocarcinoma (GI-103),  
5 weakly detected in ovarian carcinoma (GI-102); and weakly detected in colon adenocarcinoma (GI-112 and CX-1), breast carcinoma (GI-101), lung carcinoma (GI-117) and prostatic adenocarcinoma (PC3). The MTSP6 transcript was also detected in breast cancer cell line MDA-MB-231, prostate cancer cell line PC-3, but not in HT-1080 human fibrosarcoma cell line. MTSP6 also is expressed in  
10 mammary gland carcinoma cDNA (Clontech). MTSP6 also is over expressed in ovarian tumor cells.

#### MTSP7

The MTSP7 transcript was detected in lung carcinoma (A549 cell line), leukemia (K-562 cell line) and cervical carcinoma (HeLaS3 cell line). MTSP7 is  
15 believed to be expressed in lung, colon, prostate, breast, cervical and other tumors.

#### MTSP9

MTSP9 is, for example, expressed in esophageal tumor tissues, in lung carcinoma, in colorectal carcinoma, lymphoma, a cervical carcinoma (HeLaS3)  
20 and leukemia cell lines as well as in certain normal cells and tissues. MTSP9 also can be a marker for breast, prostate, cervical and colon cancer.

MTSP9 is highly expressed in the esophagus and expressed at a low level in many other tissues. The MTSP9 transcript is found in kidney (adult and fetal), spleen (adult and fetal), placenta, liver (adult and fetal), thymus,  
25 peripheral blood leukocyte, lung (adult and fetal), pancreas, lymph node, bone marrow, trachea, uterus, prostate, testes, ovary and the gland organs (mammary, adrenal, thyroid, pituitary and salivary). MTSP9 also is expressed in esophagus tumor tissues, in a lung carcinoma and, at a lower level, in a colorectal carcinoma, lymphoma, a cervical carcinoma (HeLaS3) and leukemia  
30 cell lines.

**MTSP10**

MTSP10, for example, is expressed in esophageal tumor tissues, in lung carcinoma, prostate cancers, pancreatic and breast cancers and in cell lines as well as in certain normal cells and tissues (see *e.g.*, EXAMPLES for tissue-specific expression profile). The level of activated MTSP10 can be diagnostic of prostate, uterine, lung esophagus, or colon cancer or leukemia or other cancer. The expression and/or activation of MTSP10 on or in the vicinity of a cell or in a bodily fluid in a subject can be a marker for breast, prostate, lung, colon, esophageal and other cancers.

MTSP10 transcript was detected in pancreas, lung and kidney. MTSP10 transcript was also detected in small intestine Marathon-Ready cDNA (Clontech). The MTSP10 transcript was detected in breast carcinoma (GI-101), lung carcinoma (LX-1 and GI-117), ovarian carcinoma (GI-102), and pancreatic adenocarcinoma (GI-103). The MTSP10 transcript was weakly detected in prostatic adenocarcinoma (PC3). The MTSP10 transcript was also detected in CWR22R prostate tumor grown in nude mice. No apparent signal was detected in two forms of colon adenocarcinomas (GI-112 and CX-1).

**MTSP12**

MTSP12 transcript was detected in pancreas, lung and kidney. MTSP12 transcript was also detected in small intestine Marathon-Ready cDNA (Clontech). The MTSP12 transcript was detected in breast carcinoma (GI-101), lung carcinoma (LX-1 and GI-117), ovarian carcinoma (GI-102), and pancreatic adenocarcinoma (GI-103). The MTSP12 transcript was weakly detected in prostatic adenocarcinoma (PC3). The MTSP12 transcript was also detected in CWR22R prostate tumor grown on nude mice. No apparent signal was detected in two forms of colon adenocarcinomas (GI-112 and CX-1).

**MTSP20**

MTSP20 is expressed in the lung, colon, cervical tumors and in leukemic cells. It may also be expressed in breast, ovarian, pancreatic, prostate and in other tumors. MTSP20 transcript was detected in liver, lymph node, cerebellum, pancreas, prostate, uterus, testis, glands (adrenal, thyroid and salivary), thymus, kidney and spleen. Lower transcript level was found in lung,

-70-

placenta, bladder, ovary, digestive system, circulatory system and other parts of the the brain. MTSP20 is also expressed in certain tumor cell lines including lung carcinoma (A519), colorectal carcinoma (SW480), lymphoma (Raji and Daudi), cervical carcinoma (HeLaS3) and leukemia (HL-60, K-562 and MOLT-4) cell lines.

#### MTSP22

MTSP22 is expressed in the uterine tissue, thymus, adipose tissue, and lymph node. It may also be expressed in lung, stomach, uterine, breast, ovarian, prostate and in other tumors. MTSP22 transcript was detected in some uterus tissue samples, but not in their matched tumor samples. In one of 42 uterus samples, MTSP22 is expressed in tumor and its metastatic tissues, but not in the normal tissue counterpart. MTSP22 is also expressed in some stomach tumors and lung tumors, but not in their normal tissue counterparts. MTSP22 is also expressed in the normal tissue of a pancreas matched cDNA pair. MTSP22-encoding cDNA was detected in thymus, adipose tissue, and lymph node

#### MTSP25

MTSP25 is expressed in breast, colon, uterine, ovarian, kidney, prostate, testicular cancer tissue. It may also be expressed in lung, stomach, prostate and in other tumors. MTSP25 transcript was expressed weakly in the lymph node. In the cancer profiling array analysis, MTSP25 is highly expressed in prostate samples (in normal and cancer samples). MTSP25 was highly expressed in a kidney tumor sample, but not in its normal tissue counterpart. MTSP25 was also expressed a breast cancer samples, but not in its normal tissue counterpart. MTSP25 was expressed in normal uterus samples, but not in their tumor counterparts. MTSP25 expression was also ovarian cancer samples. Among these three samples, the expression of MTSP25 was also detected in one of the matched normal tissue counterparts. MTSP25 expression was also detected in tumor samples in colon cDNA pairs.

PCR analysis revealed that MTSP25 cDNA was strongly detected in testis and mammary gland adenocarcinoma, weakly detected in brain, placenta, lung,

-71-

spleen, prostate, small intestine, colon, and leukocyte, and very weakly detected in heart, liver and pancreas.

## **2. Endotheliases**

Endotheliases are a class of cell surface proteases that are expressed on  
5 cells, particularly endothelial cells, particularly those proliferating endothelial  
cells, which are involved in a variety of proliferative processes, including  
undesirable angiogenesis associated with tumor growth and metastasis, and  
with other hyperproliferative disorders, such as restenosis, scarring, diabetic  
retinopathies, diseases and disorders of the anterior eye (see, U.S. application  
10 Serial No. 09/717,473, filed November 20, 2000, and International PCT  
application No. PCT/US00/31803).

### **Proliferative diseases**

Endotheliases are particularly useful targets for delivery of therapeutic  
agents for treatment of any disorder involving aberrant angiogenesis.

15 Endothelial cells play a key role in angiogenesis, which is  
is the generation of new blood vessels from parent microvessels. Angiogenesis  
plays a major role in the metastasis of cancer and in the pathology of a variety  
of other disorders.

Controlled and uncontrolled angiogenesis proceed in a similar manner.

20 Endothelial cells and pericytes, surrounded by a basement membrane, form  
capillary blood vessels. Angiogenesis begins with the erosion of the basement  
membrane by enzymes released by endothelial cells and leukocytes. The  
endothelial cells, which line the lumen of blood vessels, then protrude through  
the basement membrane. Angiogenic stimulants induce the endothelial cells to  
25 migrate through the eroded basement membrane. The migrating cells form a  
"sprout" off the parent blood vessel, where the endothelial cells undergo mitosis  
and proliferate. The endothelial sprouts merge with each other to form capillary  
loops, creating the new blood vessel.

-72-

### Angiogenesis, modulators and associated diseases

Angiogenesis is highly regulated by a system of angiogenic stimulators and inhibitors. Known examples of angiogenesis stimulators include certain growth factors, cytokines, proteins, peptides, carbohydrates and lipids (Norrby  
5 (1997) *APMIS* 105:417-437); Polverini (1995) *Crit. Rev. Oral. Biol. Med.* 6:230-247). A variety of endogenous and exogenous angiogenesis inhibitors are known in the art (Jackson *et al.* (1997) *FASEB* 11:457-465; Norrby (1997) *APMIS* 105:417-437); and O'Reilly (1997) *Investigational New Drugs*, 15:5-13).

Angiogenesis is essential for normal placental, embryonic, fetal and post-  
10 natal development and growth, but almost never occurs physiologically in adulthood except in very specific restricted situations. For example, angiogenesis is normally observed in wound healing, fetal and embryonal development and formation of the corpus luteum, endometrium and placenta. Angiogenesis in the adult is often associated with disease states.

15 Persistent, unregulated angiogenesis occurs in a multiplicity of disease states, tumor metastasis and abnormal growth by endothelial cells and supports the pathological damage seen in these conditions. The diverse pathological disease states in which unregulated angiogenesis is present have been grouped together as angiogenic dependent or angiogenic associated diseases.

20 The control of angiogenesis is altered in certain disease states and, in many cases, the pathological damage associated with the disease is related to uncontrolled angiogenesis (*see generally*, Norrby (1997) *APMIS* 105:417-437); and O'Reilly (1997) *Investigational New Drugs* 15:5-13). Thus, angiogenesis is involved in the manifestation or progress of various diseases, for example,  
25 various inflammatory diseases, such as rheumatoid arthritis, psoriasis, diabetic retinopathies, certain ocular disorders, including recurrence of pterygii, scarring excimer laser surgery and glaucoma filtering surgery, various disorders of the anterior eye, cardiovascular disorders, chronic inflammatory diseases, wound repair, circulatory disorders, crest syndromes, dermatological disorders (*see*,  
30 *e.g.*, U.S. Patent Nos. 5,593,990, 5,629,327 and 5,712,291) and notably cancer, including solid neoplasms and vascular tumors. Angiogenesis is essential for the growth and persistence of solid tumors and their metastases.

-73-

Repressing, eliminating or modulating this activity, should impact the etiology of these diseases and serve as a point of therapeutic intervention. In the disease state, prevention of angiogenesis could avert the damage caused by the invasion of the new microvascular system. Therapies directed at control of the

5 angiogenic processes could lead to the abrogation or mitigation of these diseases. Hence there is a need to develop therapeutics that target angiogenesis and modulate, particularly, inhibit aberrant or uncontrolled angiogenesis.

Hence conjugates that contain endotheliase substrates can be used to  
10 deliver therapeutic agents for the treatment of diseases including, but are not limited to, rheumatoid arthritis, psoriasis, diabetic retinopathies, other ocular disorders, including recurrence of pterygii, scarring from excimer laser surgery and glaucoma filtering surgery, various disorders of the anterior eye, cardiovascular disorders, autoimmune diseases, chronic inflammatory diseases,  
15 wounds, circulatory disorders, crest syndromes, restenosis, psoriasis and other dermatological disorders (see, *e.g.*, U.S. Patent Nos. 5,593,990, 5,629,327 and 5,712,291) and notably cancer, including solid neoplasms and vascular tumors.

#### Endotheliases 1 and 2

20 Exemplary of endotheliases are two different endotheliases and variant forms thereof designated endotheliase 1 and endotheliase 2 (see SEQ ID Nos. 21-27. Other members of the family can be identified by probing for genes or searching libraries for genes that have sequence identity, particularly at least 40%, 60%, 80%, 90%, 95%, 98% or greater sequence identity to the protease  
25 domain of an endotheliase identified herein, or that hybridize under conditions of high stringency to the full-length of the nucleic acid encoding a protease domain of an endotheliase provided herein, and that are expressed on endothelial cells.

Alternatively, and as a way of identifying endotheliases that can have lower sequence identity, an endotheliase can be identified by the methods, such  
30 by identifying ESTs or other nucleic acid fragments that have sequences similar to a protease and then using such fragments as probes to identify and select cDNA clones encoding full-length proteases or protease domains thereof,

-74-

identifying those that have the characteristics of transmembrane proteins, and then determining the gene expression profile to identify those that are expressed on the surface of endothelial cells. Encoded proteins that have protease activity, that include a transmembrane domain and an extracellular domain, and that are  
5 expressed in endothelial cells are endotheliases. Any method for identification of genes encoding proteins (or proteins) that encode a transmembrane protease expressed on an endothelial cell is contemplated herein.

#### **Endotheliase 1**

Exemplary of the endotheliase are endotheliase 1 and endotheliase 2.

10 These are expressed on endothelial cells. Exemplary of a full-length endotheliase 1 is one that includes the sequence of amino acids set forth in SEQ ID No. 42 (see, International PCT application No. WO 00/5006, which describes a gene it designates DESC1 that is expressed in squamous cell carcinomas and prostate tumors). As noted endotheliases are expressed on endothelial cells. A protease  
15 domain thereof is set forth in SEQ ID NO: 22.

#### **Expression profile of endotheliase 1**

To obtain information regarding the tissue distribution of endotheliase 1, the DNA insert of clone H117 was used to probe an RNA blot composed of 76 different human tissues (catalog number 7775-1; human multiple tissue  
20 expression (MTE) array; CLONTECH, Palo Alto, CA). Significant expression was observed in the esophagus, with minor expression levels in the stomach, salivary gland, pancreas, prostate, bladder, trachea and uterus. Northern analysis using RNA blots (catalog numbers 7765-1 & 7782-1; human muscle and digestive system multiple tissue northern (MTN) blots; CLONTECH) confirmed that the  
25 expression was restricted to the esophagus. Two transcripts (approximately 1.7 and 2 kb) were detected in the esophagus. Endotheliase 1 also is expressed in umbilical vein endothelial cells, PC3 and LnCAP cells.

#### **Endotheliase 2 and nucleic acids encoding endotheliase 2**

Two splice variant forms of endotheliase 2 designated endotheliase 2-S  
30 and endotheliase 2-L are exemplified herein (see SEQ ID Nos. 23-26). The open reading frame of the nucleic acid encoding endotheliase 2-S (SEQ ID No. 23) is composed of 1,689 bp, which translates to a 562-amino acid protein (SEQ ID

-75-

No. 24), while the ORF of endotheliase 2-L is composed of 2,067 bp (SEQ ID No. 25), which translates to a 688-amino acid protein (SEQ ID No. 26).

The nucleic acid encoding the protease domain of endotheliase 2-S is composed of 729 bp which translates to a 242-amino acid protein (amino acids 321-562 of SEQ ID Nos. 23 and 24), while that of endotheliase 2-L is composed of 1,107 bp, which translates to a 368-amino acid protein (amino acids 321-688 of SEQ ID Nos. 25 and 26).

#### **Endotheliase-2 proteins**

Any and all of the above-noted endotheliases and/or protease domains thereof, such as those that include the sequences of amino acids in SEQ ID Nos. 22, 24, 26 and 27 or are encoded by nucleic acid that hybridize thereto under the conditions as described above are contemplated for use in the methods herein. Also contemplated herein are proteins that include amino acid sequence changes, such as those set forth in Table 1 above, and retain protease activity.

#### **Gene expression profile and transcript size of endotheliase 2 in normal and tumor tissues**

In addition to expression in endothelial cells, endotheliase 2 is expressed in placenta, pancreas, thyroid gland, liver and lung tissues. It also is expressed at lower levels in mammary gland, salivary gland, kidney, trachea, esophagus, appendix, heart and fetal lung. Endotheliase 2 also is expressed in several tumor cell lines and, hence, in certain tumors, including lung and colon, including breast carcinoma, lung carcinomas, colon adenocarcinomas, pancreatic adenocarcinoma (GI-103), and ovarian carcinoma. It has also been detected in prostate and fibrosarcoma cell lines.

#### **C. Conjugates**

Conjugates that are substrates for proteases on the surfaces of cells, particularly serine proteases, including type II membrane-bound serine proteases, and endotheliases are provided. Any cell surface protease, including cell-associated or localized proteases, is contemplated herein. Generally proteases expressed at high levels in active forms in essential tissues are not ideal target candidates. The proteases include those that are expressed on relatively limited numbers of cells or that are expressed at high levels in cells, such as tumor cells



-76-

and endothelial cells and immune cells, that are involved in disease states or are present in diseases states in the locale of cells involved in the disease states.

For example, endothelial cells by virtue of their role in angiogenesis are involved in numerous proliferative disorders; immune cells are involved in many disease

5 processes including cancers and diseases and inflammatory disorders. Other cell surface proteases are expressed at higher levels in certain tumors than in normal cells. Whether or not such proteases have a role in the disorder their higher expression in cells involved in a disease state is sufficient for use for targeting therapeutic agents in the conjugates provided herein.

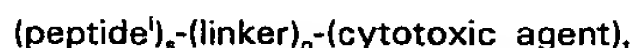
10 The conjugates, which contain a therapeutic agent, such as a cytotoxic agent, is activated upon cleavage by a cell surface protease, including cell-associated and cell-localized proteases. Exemplary of such proteases are the MTSPs, such as, but not limited to, MTSP1, MTSP3, MTsP4, MTSP6, MTSP7, MTSP9, MTSP10, MTSP12, MTSP20, MTSP22, MTSP25, urokinases and  
15 endotheliases. Hence, the conjugates targeted to such proteases are prodrugs in that the therapeutic agent is inactive as administered and is ultimately activated in the vicinity of the targeted cell or tissue. Although cell surface proteases, such as transmembrane proteases, are the intended targets, any released, shed or soluble forms of the proteases and others also can be targeted.

20 Thus, the conjugates, which contain a therapeutic agent, such as a cytotoxic agent, are substantially inactive prior to action by a cell surface protease, a peptidic moiety that is a substrate for a targeted cell surface protease (*i.e.*, a peptidic substrate), and, optionally, a linker. The therapeutic agents in the conjugates are activated upon cleavage of the peptidic substrate of  
25 the conjugate by a cell surface protease. The therapeutic agents, such as cytotoxic agents, are released as the free yagent, or, alternatively, are released coupled to the portion of the peptidic substrate (P1-P2-P3-etc. (*i.e.*, the N-terminus) or P1'-P2'-etc. (*i.e.*, the C-terminus) that the agents were linked to in the conjugate, optionally via a linker. The cytotoxic agents, in these forms, are  
30 released in the vicinity of cells that express the proteases. Activation is effected, in certain embodiments, because the therapeutic agent, such as cytotoxic agent, following action of the cell surface protease, can cross the cell

-77-

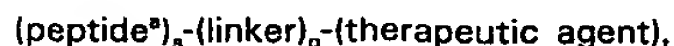
membrane or otherwise interact with the cell or tissue and exhibit therapeutic activity. In other embodiments, any remaining peptidic moieties or amino acids can be cleaved from therapeutic agent to render it active. The conjugates act as prodrugs because the therapeutic agents when conjugated are substantially  
 5 inactive. Upon cleavage by the targeted protease, the therapeutic agent is released either in active form or in a form that is activated by the targeted cell, tissue or surrounding environment.

In one exemplary embodiment, the targeted agent is a cytotoxic agent and the conjugates for use in the methods and compositions provided herein  
 10 have the formula:



or a derivative thereof, where peptide<sup>i</sup> is a peptidic substrate for a cell surface protease or a released, shed or otherwise unbound membrane protease, such as an MTSP; s is greater than or equal to 1, or is 1 to 6, or is 1 or 2, or is 1; linker  
 15 is any linker; q is greater than or equal to 0, or is 0 to 4, or is 0 or 1; the cytotoxic agent is an anti-tumor, anti-cancer or anti mitotic agent, including anti-angiogenic agents; and t is 1 or more, or is 1 or 2. In these conjugates, the cytotoxic agent is covalently attached, optionally via a linker, to either the C-terminus or the N-terminus of the peptidic substrate. In embodiments where the  
 20 therapeutic agent, such as a cytotoxic agent, is attached to the C-terminus of the peptidic substrate, the N-terminus optionally is capped. N-Terminal caps for use herein include, but are not limited to, acyl, sulfonyl and carbamoyl groups. In embodiments where the therapeutic agent is attached to the N-terminus of the peptidic substrate, the C-terminus is a carboxamide derivative.

25 In certain embodiments, peptide<sup>i</sup> is a peptidic substrate for a cell surface protease or a soluble MTSP whereby, upon action of the protease, the conjugate, which is substantially inactive, is cleaved at the P1-P1' bond to release a compound of the formula:



30 or a derivative thereof, that exhibits therapeutic activity, such as cytotoxic activity *in vitro* and *in vivo*. In these compounds, peptide<sup>a</sup> is a truncated version of peptide<sup>i</sup> resulting from cleavage at the P1-P1' bond.

-78-

In another embodiment, the conjugates for use in the methods and compositions provided herein possess two therapeutic agents, such as cytotoxic agents, which are the same or different, linked to the C-terminus and the N-terminus, respectively, optionally via linkers linker<sup>1</sup> and linker<sup>2</sup>, of a peptidic substrate for cell surface protease or a soluble MTSP. In this embodiment, the

conjugates have the formula:  
 (therapeutic agent<sup>1</sup>)<sub>x</sub>-(linker<sup>1</sup>)<sub>w</sub>-(peptide<sup>1</sup>)<sub>s</sub>-(linker<sup>2</sup>)<sub>q</sub>-(therapeutic agent<sup>2</sup>)<sub>t</sub>,  
 or a derivative thereof, where peptide<sup>1</sup> is a peptidic substrate for a cell surface protease, or a soluble MTSP; s is greater than or equal to 1, or is 1 to 6, or is 1  
 or 2, or is 1; linker<sup>1</sup> and linker<sup>2</sup> are each independently any linker and are the  
 same or different; q and w are each independently greater than or equal to 0, or  
 are 0 to 4, or are 0 or 1; the therapeutic agents, which are the same or  
 different, are anti-tumor, anti-cancer or anti mitotic agents; and t and x are each  
 independently 1 or more, or are 1 or 2.

In these embodiments, peptide<sup>1</sup> is a peptidic substrate for a cell surface protease or a soluble MTSP whereby, upon action of the protease, the conjugate, which is substantially inactive, is cleaved at a point on the peptidic chain to release two compounds of the formulae:

(therapeutic agent<sup>1</sup>)<sub>x</sub>-(linker<sup>1</sup>)<sub>w</sub>-(peptide<sup>a1</sup>)<sub>s</sub>; and

(peptide<sup>a2</sup>)<sub>s</sub>-(linker<sup>2</sup>)<sub>q</sub>-(therapeutic agent<sup>2</sup>)<sub>t</sub>,

or derivatives thereof. The released therapeutic agents are active or are further activated by the cell, tissue or surrounding environment. In these compounds, peptide<sup>a1</sup> and peptide<sup>a2</sup> are N-terminal and C-terminal truncated portions, respectively, of peptide<sup>1</sup> resulting from cleavage at the P1-P1' bond.

In one embodiment, the conjugates for use in the compositions and methods provided herein have formula I:

$X^n-(P6)_m-(P5)_p-(P4)_i-(P3)_j-(P2)_l-P1-(P1')_u-(P2')_k-(P3')_r-(L)_n-Z$

or a derivative thereof, where Z is a therapeutic agent; L is a linker; l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

-79-

u, k and r are selected as follows:

u is 0 or 1; when u is 0, k and r are 0; when u is 1, k is 0 or 1; when k is 0, r is 0; when k is 1, r is 0 or 1;

n is 0 or 1;  $X^n$  is hydrogen, or an acyl, sulfonyl or carbamoyl cap; and P6 to P3' are amino acid residues, as defined below. In this embodiment, the P6 to P3' residues are linked by peptide bonds or peptide bond surrogates. Thus, the P6 to P3' portion of the conjugate is a peptidic substrate, as defined herein.

In another embodiment, the conjugates for use in the compositions and methods provided herein have formula II:

10  $Z-(L)_n-(P6)_m-(P5)_p-(P4)_i-(P3)_j-(P2)_l-P1-(P1')_u-(P2')_k-(P3')_r-X^c$

or a derivative thereof, where Z is a therapeutic agent; L is a linker; l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k and r are selected as follows:

u is 0 or 1; when u is 0, k and r are 0; when u is 1, k is 0 or 1; when k is 0, r is 0; when k is 1, r is 0 or 1;

n is 0 or 1;  $X^c$ , together with the carbonyl group of the amino acid residue to which it is attached, forms a carboxylic acid or a carboxamide group; and P6 to P3' are amino acid residues, as defined below. In this embodiment, the P6 to P3' residues are linked by peptide bonds or peptide bond surrogates. Thus, the P6 to P3' portion of the conjugate is a peptidic substrate, as defined herein.

25 In a further embodiment, the conjugates for use in the compositions and methods provided herein have formula III:

$Z^1-(L^1)_n-(P6)_m-(P5)_p-(P4)_i-(P3)_j-(P2)_l-P1-(P1')_u-(P2')_k-(P3')_r-(L^2)_v-Z^2$

or a derivative thereof, where  $Z^1$  and  $Z^2$  are each therapeutic agents and are the same or different;  $L^1$  and  $L^2$  are each linkers and are the same or different; l, j, i,

30 p and m are selected as follows:

-80-

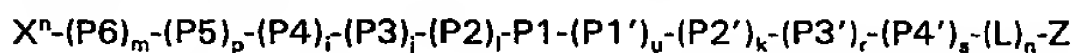
l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k and r are selected as follows:

5 u is 0 or 1; when u is 0, k and r are 0; when u is 1, k is 0 or 1; when k is 0, r is 0; when k is 1, r is 0 or 1;

n and v are each independently 0 or 1; and P6 to P3' are amino acid residues, as defined below. In this embodiment, the P6 to P3' residues are linked by peptide bonds or peptide bond surrogates. Thus, the P6 to P3' portion  
10 of the conjugate is a peptidic substrate, as defined herein.

In another embodiment, the conjugates for use in the compositions and methods provided herein have formula IV:



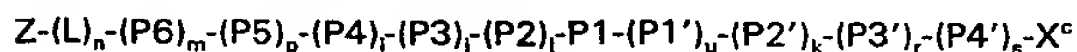
or a derivative thereof, where Z is a therapeutic agent; L is a linker; l, j, i, p and  
15 m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k, r and s are selected as follows:

20 u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1; when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0; when r is 1, s is 0 or 1;

n is 0 or 1; X<sup>n</sup> is hydrogen, or an acyl, sulfonyl or carbamoyl cap; and P6 to P4' are amino acid residues, as defined below. In this embodiment, the P6 to P4' residues are linked by peptide bonds or peptide bond surrogates. Thus, the  
25 P6 to P4' portion of the conjugate is a peptidic substrate, as defined herein. In another embodiment, the conjugates for use in the compositions and methods provided herein have formula V:



30 or a derivative thereof, where Z is a therapeutic agent; L is a linker; l, j, i, p and m are selected as follows:

-81-

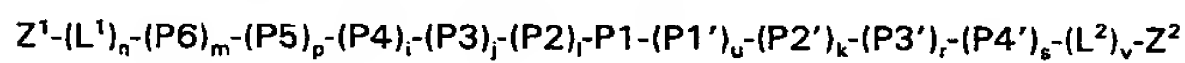
l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k, r and s are selected as follows:

- 5 u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1; when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0; when r is 1, s is 0 or 1;

n is 0 or 1; X<sup>c</sup>, together with the carbonyl group of the amino acid residue to which it is attached, forms a carboxylic acid or a carboxamide group; and P6 to P4' are amino acid residues, as defined below. In this embodiment, the P6 to P4' residues are linked by peptide bonds or peptide bond surrogates. Thus, the P6 to P4' portion of the conjugate is a peptidic substrate, as defined herein.

- 15 In a further embodiment, the conjugates for use in the compositions and methods provided herein have formula VI:



or a derivative thereof, where Z<sup>1</sup> and Z<sup>2</sup> are each therapeutic agents and are the same or different; L<sup>1</sup> and L<sup>2</sup> are each linkers and are the same or different; l, j, i, p and m are selected as follows:

- 20 l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k, r and s are selected as follows:

- 25 u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1; when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0; when r is 1, s is 0 or 1;

- 30 n and v are each independently 0 or 1; and P6 to P4' are amino acid residues, as defined below. In this embodiment, the P6 to P4' residues are linked by peptide bonds or peptide bond surrogates. Thus, the P6 to P4' portion of the conjugate is a peptidic substrate, as defined herein.

Exemplary peptidic substrates, therapeutic agents, linkers and exemplary conjugates of formulae I-VI are described in further detail below. It is intended

-82-

herein that conjugates resulting from all combinations and/or permutations of the groups recited below for the variables of formulae I-VI are encompassed within the instant disclosure.

#### 1. Peptidic Substrates

5       The peptidic substrates contemplated for use in the conjugates are substrates for the targeted cell surface protease or a soluble, shed or released form thereof, and contain a sufficient number of amino acid residues to render any therapeutic agent in the conjugate substantially inactive. In the exemplary embodiment where the therapeutic agent is, for example, doxorubicin, the  
10       conjugate is substantially inactive by virtue of the inability of the conjugated therapeutic agent to cross the cell membrane. In certain embodiments, the peptidic substrate contains at least 1, 2, 3, 4 or 5 amino acid residues, and can contain up to nine or ten residues. Longer peptidic substrates can be used in the conjugates as long as upon cleavage, the resulting therapeutic agent or  
15       therapeutic agent-amino acid or -peptidic moiety conjugate exhibits the desired therapeutic effect *in vivo* and *in vitro*.

Hence, exemplary peptidic substrates for use in the conjugates provided herein possess at least one amino acid (P1), two amino acids (P1-P1'), three amino acids (P2-P1-P1') and typically contain four, five or six amino acid  
20       residues (P3-P2-P1-P1', P4-P3-P2-P1-P1' or P4-P3-P2-P1-P1'-P2'), where the P1-P1' bond is the site of cleavage of cell surface protease, or a soluble, shed or released form thereof, including, but not limited to, a cell surface protease, such as a serine protease, including, for example, but not limited to, uPA bound to its receptor, MTSPs and endotheliases. The peptidic substrates optionally further  
25       possess a P5, P6 or P3' amino acid residue, and, in certain embodiments, possess P7, P8, P9, P10, P4', P5', P6' residues. Thus, the peptidic substrates for use in the conjugates provided herein are penta-, hexa-, hepta-, octa- and nona-peptidic substrates, and can contain 10, 11, 12, 13, 14, 15 or more residues as long as, upon cleavage of the conjugate by the protease, the  
30       resulting therapeutic agent or therapeutic agent-amino acid or -peptidic moiety conjugate exhibits the desired therapeutic effect *in vivo* and *in vitro*.

-83-

The peptidic substrates are conjugated to the therapeutic agent (or to a linker to which the therapeutic agent is linked) via the C-terminal residue (*i.e.*, P1', P2' or P3'), or the N-terminal residue (*i.e.*, P6, P5 or P4), or optionally an internal residue. The peptidic substrates, for example, can be straight chains,  
 5 but can be cyclized or include cyclized portions.

In embodiments where the conjugation is via the C-terminus of the peptidic substrate, the peptidic substrate optionally possesses a cap, such as an acyl or carbamoyl cap at the N-terminus. In embodiments where conjugation is via the N-terminus of the peptidic substrate, the peptidic substrate further  
 10 possess a terminal group, such as a carboxamide group, at the C-terminus.

The conjugates can contain a plurality of peptidic substrates and a plurality of therapeutic agents. For example, in conjugates that contain two therapeutic agents, which are the same or different, conjugation to the therapeutic agent(s) or linker linked thereto can be via the C-terminal and N-  
 15 terminal residues of the peptidic substrate.

The methods described for selection of substrates above can be used to design suitable substrates. In addition, substrates can be designed based upon known specificities of other proteases. For example, the specificities of trypsin-like and trypsin family members can aid in design of possible substrates. The  
 20 following summarizes substrate preferences for particular serine proteases (see, *e.g.*, Harris *et al.* (2000) *PNAS* 97(14):7754-7759).

PROTEASE	EXEMPLARY P1 RESIDUE(S)	EXEMPLARY P2 RESIDUE(S)	EXEMPLARY P3 RESIDUE(S)
Chymotrypsin	Tyr, Phe, Trp	--	--
Trypsin	Arg, Lys	--	--
25 Thrombin	Arg, Lys	Phe	Thr, Trp
Plasmin	Lys, Arg	Trp, Tyr, Met	Gln
Granzyme B	Asp	--	--
Human Neutrophil Elastase	Ala, Val, Ile	--	--
30 Tissue Plasminogen Factor	Arg	Ser, Gly, Ala	Met, Tyr
Urokinase	Arg	Ser, Ala	Thr, Ser
Factor Xa	Arg	Gly	--



-84-

Typical protocols for preparation of the conjugates can include the steps of: 1) identification of a targeted protease; 2) expression and assay development; 3) substrate selection, such as, for example, by testing chromogenic or fluorogenic substrates to identify those cleaved by a selected target protease, by use of substrate phage display to identify peptidic substrates cleaved by a targeted protease, by use of a natural protein or peptide substrate or a natural inhibitor of the protease, and by use of combinatorial libraries to identify substrates cleaved by a targeted protease; 4) synthesis of conjugates containing the identified substrate; and 5) biological evaluation thereof, including, but not limited to, *in vitro* assays, cell culture assays, biological assays, and *in vivo* animal models (see, e.g., EXAMPLE 10).

A conjugate can be designed by any methods known to those of skill in the art. The following provides an exemplary protocol. First, a series of commercially available chromogenic and fluorogenic peptidic substrates can be tested for cleavage by the protease of interest (see Examples for lists of exemplary chromogenic and fluorogenic substrates and the table below). The peptidic portion of these substrates occupies the unprimed binding sites of the protease while the reporter group is located on the primed side of the scissile bond. Effective conjugates can then be designed based on the structure of the substrates that are efficiently cleaved by the protease.

The peptidic portion of these efficiently cleaved substrates can be used as the unprimed region of the conjugate, and Ser-therapeutic agent, such as a cytotoxic agent (e.g., doxorubicin), Ser-Leu-therapeutic agent or Ser-Ser-Leu-therapeutic agent can be used as the primed region of the conjugate. Cleavage of these conjugate prodrugs releases either Ser-therapeutic agent, Ser-Leu-therapeutic agent or Ser-Ser-Leu-therapeutic agent compounds. In another embodiment, the Ser in the released Ser-therapeutic agent may be replaced by other amino acid residues including, but not limited to, Ala, hSer, Abu, Thr, Met, nLeu and Val. In another embodiment, such as when the therapeutic agent is doxorubicin, the amino acid residue conjugated to the therapeutic agent possesses a hydrophobic side chain. Such amino acid residues include, but are not limited to, Leu, Abu, nLeu, nVal, CHA, hCHA, (hex)Gly, (allyl)Gly,

-85-

(propargyl)Gly and (cyclopropyl)Ala. In another embodiment, such as when the therapeutic agent is taxol, the amino acid residue conjugated to the therapeutic agent possesses a side chain that is not sterically bulky. Such amino acid residues include, but are not limited to, Gly and Ala. The resulting P1'-

5 therapeutic agent, P1'-P2'-therapeutic agent, or P1'-P2'-P3'-therapeutic agent compound can be further processed *in vivo* into active therapeutic agents.

Another approach to designing a conjugate prodrug for a protease substrate is to use substrate phage display to elucidate optimal subsite occupancy for the protease. The resulting information can then be used to  
10 design the peptidic, unprimed portion of the conjugate. As described above, the primed region of the conjugate can be fixed as Ser-therapeutic agent, Ser-Leu-therapeutic agent or Ser-Ser-Leu-therapeutic agent.

A third approach to design an effective prodrug conjugate involves the use of combinatorial fluorogenic substrate libraries to determine optimal residues  
15 for the unprimed region of a protease substrate. These selected sequences can then be used as the unprimed portion of the conjugate prodrug and, and Ser-therapeutic agent, (e.g., doxorubicin), Ser-Leu-therapeutic agent or Ser-Ser-Leu-therapeutic agent can be used as the primed region of the conjugate.

These methods have been used in the design of the peptidic substrate portion of  
20 the conjugates provided herein. For example, sequences including GSGR (and related sequences such as TGR, SGR, extended variants and others herein) were based on or derived from substrate phage display experiments using u-PA as the target protease. Many matriptase conjugates, such as (R/K)-X-S-R and X-(R/K)-S-R, and related sequences as provided herein, were based on data from  
25 combinatorial libraries. In other embodiments, sequence sequences in natural substrates or natural inhibitors of a protease target, such as uPA, including VSAR, PGR (from P3-P1 of plasminogen) and related sequences, were used in design of u-PA-targeted conjugates. In other embodiments, sequences from chromogenic substrates, such as D-HHT-Gly-Arg, and related sequences, were  
30 used for design of ET-1-targeted conjugates.

#### Chromogenic/fluorogenic substrates

Enzyme	Substrate	Structure
--------	-----------	-----------

-86-

5	MTSP1	spectrozyme t-PA	CH <sub>3</sub> SO <sub>2</sub> -D-HHT-Gly-Arg-pNA.AcOH
	MTSP1	S 2765	N- $\alpha$ -Z-D-Arg-Gly-Arg-pNA.2HCl
	MTSP3	Spectrozyme fXIIa	H-D-CHT-Gly-Arg-pNA.2AcOH
	MTSP4 <sup>a</sup>	Spec PL	H-D-Nle-HHT-Lys-pNA.2AcOH
	MTSP5	S 2765	N- $\alpha$ -Z-D-Arg-Gly-Arg-pNA.2HCl
10	MTSP6	spectrozyme t-PA	CH <sub>3</sub> SO <sub>2</sub> -D-HHT-Gly-Arg-pNA.AcOH
	MTSP7	S 2366	pyroGlu-Pro-Arg-pNA.HCl
	MTSP9	Pefachrome fVIIa	CH <sub>3</sub> SO <sub>2</sub> -D-CHA-But-Arg-pNA
	MTSP10	spectrozyme t-PA	CH <sub>3</sub> SO <sub>2</sub> -D-HHT-Gly-Arg-pNA.AcOH
	MTSP22	S 2366	pyroGlu-Pro-Arg-pNA.HCl
15	ET-1	spectrozyme t-PA	CH <sub>3</sub> SO <sub>2</sub> -D-HHT-Gly-Arg-pNA.AcOH
	ET-2	S 2765	N- $\alpha$ -Z-D-Arg-Gly-Arg-pNA.2HCl
	u-PA	S-2444	pyroGlu-Gly-Arg-pNA.HCl

- 15 <sup>a</sup> coupled assay, activation of plasminogen in the presence of Spec PL

Briefly, for a coupled assay, the ability of the protease to activate an enzyme, such as plasminogen or trypsinogen is tested. To perform these assays, a protease is incubated with a zymogen, such as plasminogen or trypsinogen, in the presence of a labelled known substrate, such as lys-plasminogen or Spec PL (for plasmin), for the zymogen. If protease activates the zymogen, the activated enzyme, such as plasmin and trypsin, will degrade the substrate, thereby changing the spectral properties of the substrate.

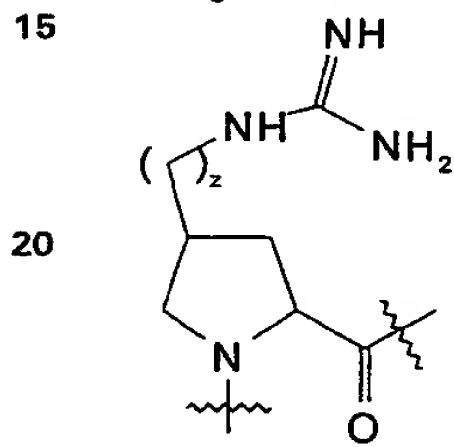
## 25 Exemplary peptidic substrates

The following description provides exemplary peptidic substrates for cleavage by proteases, such as MTSP1 (or matriptase), endotheliase 1 and urokinase, and a general discussion of properties of the residues. In a similar manner, peptidic substrates for cleavage by other cell surface proteases, or a soluble, shed or released form thereof, can be similarly designed by identifying peptidic substrates for the selected protease and then preparing conjugates that contain such peptidic substrates.

-87-

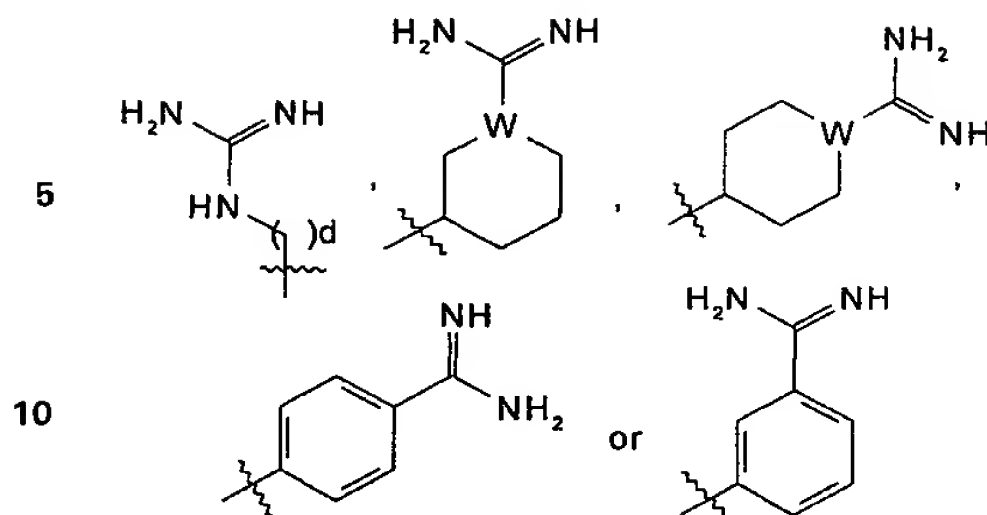
a. The P1 Residue

Amino acid residues for use at the P1 position of the peptidic substrates for use in the conjugates provided herein include Arg, Arg surrogates and Lys. Arg surrogates include unnatural amino acids that possess a group or moiety that functions in substantially the same way as the naturally occurring side chain of arginine to achieve substantially the same result (*e.g.*, acting as the P1 residue in a substrate for a MTSP1, urokinase or endotheliase). Arg surrogates include, but are not limited to,  $\alpha$ -amino acids that possess as the side chain any of the following: the side chain of homoarginine; guanidinoaminopropyl; guanidinoaminoethyl; (Me)<sub>2</sub>arginine side chain; (Et)<sub>2</sub>arginine side chain; (4-aminomethyl)phenylmethyl; 4-amidinophenylmethyl; 4-guanidinophenylmethyl; or the Arg surrogate is a conformationally constrained arginine analog such as:



where z is 0 or 1 (see, *e.g.*, Webb *et al.* (1991) *J. Org. Chem.* 56:3009); or the side chain is a conformationally constrained arginine side chain analog such as:

-88-



15 where d is an integer from 0 to 5, or 1 to 3; and W is N or CH; or a mono- or di-substituted N-alkyl derivative of the above groups, where alkyl is, in certain embodiments, lower alkyl, such as, for example, methyl.

In certain embodiments herein, the P1 residue is Arg.

#### 20 b. The P2 Residue

In the conjugates provided herein, the P2 residue is selected from Phe, Ser, Gly, Ala, Ser(OMe), hSer, 1-methylHis, 3-methylHis, His, nVal, nLeu, Abu, (hS)Gly, Thr, Aib, CHA and Tyr. In another embodiment, the P2 residue is selected from Phe, Ser, Gly and Ala. In certain  
 25 embodiments herein, the P2 residue is Ser or Ala. In another embodiment, the P2 residue is Gly or Ala.

#### c. The P3 Residue

Amino acid residues for use at the P3 position of the conjugates provided herein include Arg, Lys, Gln, Quat, Arg surrogates, Ser, Thr,  
 30 hSer, dSer, Pro, (hS)Gly, Tyr, 4,4-dimethylThr, Asn, Met(O<sub>2</sub>), Quat<sup>2</sup>, Quat<sup>3</sup>, Quat<sup>4</sup> and Quat<sup>5</sup>. In another embodiment, the P3 residue is selected from Arg, Lys, Gln, Quat and Arg surrogates. Arg surrogates include those described above for the P1 residue.

In certain embodiments, the P3 residue is Gln or Ser.

**d. The P4 Residue**

In the conjugates provided for use in the compositions and methods provided herein, the P4 residue is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe, Val, N,N-dimethylGly,  $\beta$ -Ala, Cys(Me), Gln, t-butylGly and nVal. In another embodiment, the P4 residue is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val. In further embodiments, the P4 residue is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Phe or Val. In certain embodiments herein, the P4 residue is Arg or Gly.

**10 e. The P5 and P6 Residues**

In certain embodiments herein, the peptidic substrates used in the conjugates contain a P5 and, optionally, a P6 residue. P5 residues include Ile, Arg and Arg surrogates. In another embodiment, P5 residues include Arg and Arg surrogates. Arg surrogates include those described above for the P1 residue. P6 residues include, for example, Leu, Val and Arg. In another embodiment, P6 residues include, for example, Leu.

**f. The P1' Residue**

The P1' residue of the conjugates provided herein is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, Thr or hSer. In another embodiment, the P1' residue of the conjugates provided herein is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl. In another embodiment, the P1' residue is Ser, Ala, hSer, Abu, Thr, Met, nLeu or Val. In another embodiment, the P1' residue is Gly or Ala. In another embodiment, the P1' residue is Ser, Ala or Gly. In another embodiment, the P1' residue is Leu, Abu, nLeu, nVal, CHA, hCHA, (hex)Gly, (allyl)Gly, (propargyl)Gly or (cyclopropyl)Ala. In certain embodiments herein, the P1' residue is Ala, Ser, Gly, Ile or d-Ile.

-90-

**g. The P2' Residue**

In certain embodiments herein, the conjugates provided herein possess a P2' residue. P2' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, hCHA, CHA, hexylGly, allylGly and Phe. In another embodiment, P2' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl. In another embodiment, the P2' residue is Ser, hSer, Abu, nLeu, nVal, CHA, hCHA, (allyl)Gly or (hexyl)Gly. In another embodiment, the P2' residue is Gly or Ala. In another embodiment, the P2' residue is Leu, Abu, nLeu, nVal, CHA, hCHA, (hex)Gly, (allyl)Gly, (propargyl)Gly or (cyclopropyl)Ala. In further embodiments, the P2' residues are Ala, Gly, Ile or d-Ile.

**h. The P3' Residue**

In other embodiments herein, the peptidic substrates used in the conjugates provided herein include a P3' residue. P3' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly. In another embodiment, the P2' residue is Ser, hSer, Abu, nLeu, nVal, CHA, hCHA, (allyl)Gly or (hexyl)Gly. In another embodiment, P3' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl. In another embodiment, the P3' residue is Gly or Ala. In another embodiment, the P3' residue is Leu, Abu, nLeu, nVal, CHA, hCHA, (hex)Gly, (allyl)Gly, (propargyl)Gly or (cyclopropyl)Ala.

**i. The P4' Residue**

In other embodiments herein, the peptidic substrates used in the conjugates provided herein include a P4' residue. P4' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, nLeu, Val,

-91-

nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly. In another embodiment, P4' residues for use herein include, but are not limited to, Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl. In another embodiment, the P4' residue is Gly or Ala. In another embodiment, the P4' residue is Leu, Abu, nLeu, nVal, CHA, hCHA, (hex)Gly, (allyl)Gly, (propargyl)Gly or (cyclopropyl)Ala. In another embodiment, the P4' residue is Leu.

j. Caps

1) X<sup>n</sup> (the N-terminal Cap)

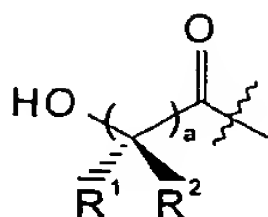
- 10 In embodiments herein where the therapeutic agent is conjugated to the C-terminus of the peptidic substrate (*i.e.*, where the conjugate has formula I), the N-terminus of the peptidic substrate optionally is capped with an acyl, sulfonyl or carbamoyl derivative. The cap is chosen, in certain embodiments, to increase the hydrophilicity of the conjugate. In
- 15 embodiments where the peptidic substrate-therapeutic agent conjugate is sufficiently hydrophilic so as not to require further hydrophilicity, a non-hydrophilic N-terminal cap, such as an acetyl group, can be used. In embodiments where increased hydrophilicity is desired, the N-terminal amino acid is modified with a hydrophilic blocking group. Such blocking
- 20 groups are chosen based upon the presence of hydrophilic functionality. Such blocking of the terminal amino group can also reduce or eliminate the enzymatic degradation of such peptidyl therapeutic agents by the action of exogenous amino peptidases which are present in the blood plasma of warm blooded animals.
- 25 N-Terminal blocking groups that increase the hydrophilicity of the conjugates and therefore increase the aqueous solubility of the conjugates include, but are not limited to, hydroxylated alkanoyl, polyhydroxylated alkanoyl, polyethylene glycol, glycosylates, sugars and crown ethers.



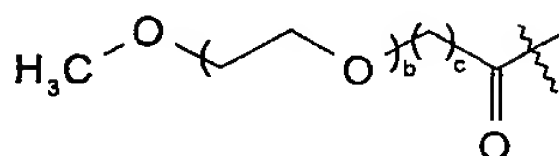
-92-

In certain embodiments herein, the N-terminal blocking group is one of the following:

a)



or b)



where  $R^1$  and  $R^2$  are selected from (i) or (ii) as follows:

(i)  $R^1$  and  $R^2$  are each independently:

a) hydrogen;

b) unsubstituted or substituted aryl, unsubstituted or substituted heterocyclyl,  $C_3$ - $C_{10}$  cycloalkyl,  $C_2$ - $C_6$  alkenyl,  $C_2$ - $C_6$  alkynyl, halogen,  $C_1$ - $C_6$  perfluoroalkyl,  $R^4O$ -,  $R^3C(O)NR^3$ -,  $(R^3)_2NC(O)$ -,  $(R^3)_2N-C(NR^3)$ -,  $R^4S(O)_eNH$ -, -CN, -NO<sub>2</sub>,  $R^3C(O)$ -, -N<sub>3</sub>, -N( $R^3$ )<sub>2</sub>, or  $R^4OC(O)NR^3$ -;

c) unsubstituted  $C_1$ - $C_6$  alkyl;

d) substituted  $C_1$ - $C_6$  alkyl wherein the substituent on the substituted  $C_1$ - $C_6$  alkyl is selected from unsubstituted or substituted aryl, unsubstituted or substituted heterocyclyl,  $C_3$ - $C_{10}$  cycloalkyl,  $C_2$ - $C_6$  alkenyl,  $C_2$ - $C_6$  alkynyl,  $R^3O$ -,  $R^4S(O)_eNH$ -,  $R^3C(O)NR^3$ -,  $(R^3)_2NC(O)$ -,  $(R^3)_2N-C(NR^3)$ -, -CN,  $R^3C(O)$ -, -N<sub>3</sub>, -N( $R^3$ )<sub>2</sub>, and  $R^4OC(O)NR^3$ -; or

-93-

(ii)  $R^1$  and  $R^2$  are combined to form  $-(CH_2)_f-$  where one of the carbon atoms optionally is replaced by a moiety selected from:  $-O-$ ,  $-S(O)_e-$ ,  $-NC(O)-$ ,  $-NH-$  and  $-N(COR^4)-$ ;

$R^3$  is selected from: hydrogen, unsubstituted or substituted aryl, unsubstituted or substituted heterocyclyl,  $C_1-C_6$  alkyl and  $C_3-C_{10}$  cycloalkyl;

$R^4$  is selected from: unsubstituted or substituted aryl, unsubstituted or substituted heterocyclyl,  $C_1-C_6$  alkyl and  $C_3-C_{10}$  cycloalkyl;

$e$  is 0, 1 or 2;

$a$  is 1, 2, 3 or 4;

$b$  is zero or an integer between 1 and 100; and

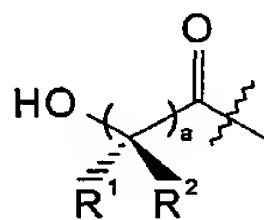
$c$  is 0 to 10, provided that if  $b$  is zero,  $c$  is 1 to 10; and

$f$  is 3, 4 or 5.

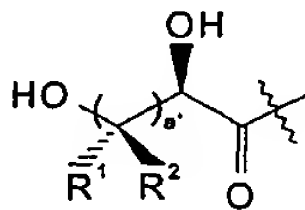
In certain embodiments,  $R^1$  and  $R^2$  are each independently hydrogen, OH,  $C_1-C_6$  alkyl,  $C_1-C_6$  alkoxy,  $C_1-C_6$  aralkyl or aryl. In these embodiments,  $a$  is 1, 2, 3 or 4;  $b$  is 0 or an integer between 1 and 100; and  $c$  is 0 to 10, provided that if  $b$  is 0,  $c$  is 1 to 10.

In another embodiment, the N-terminal cap ( $X^n$ ) is hydrogen, or (i), (ii), (iii) or (iv) as follows:

(i)

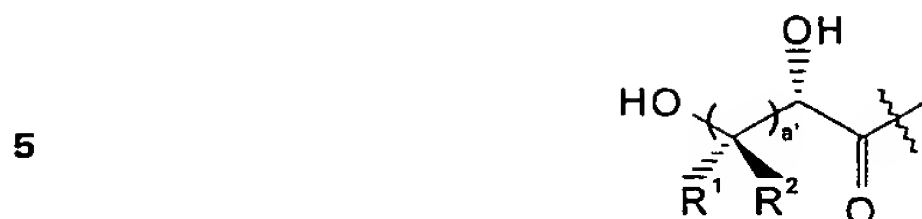


or (ii)

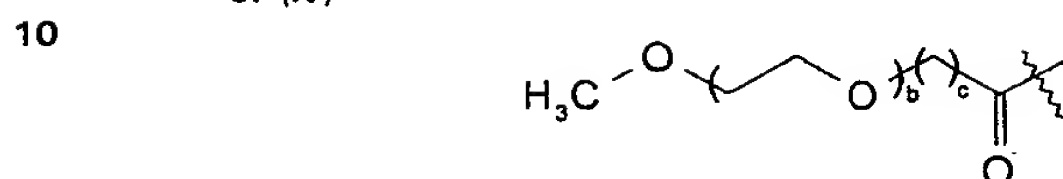


-94-

or (iii)



or (iv)



- 15 where  $R^1$  and  $R^2$  are each independently hydrogen,  $C_1$ - $C_6$  alkyl and aryl;  $a$  is 1, 2, 3 or 4;  $a'$  is 0, 1, 2 or 3;  $b$  is 0 or an integer between 1 and 14; and  $c$  is 0 or 1, provided that if  $b$  is 0,  $c$  is 1.

- In another embodiment,  $X^n$  is  $R^{30}O-C(O)-$ ,  $R^{31}R^{32}N-C(O)-$ ,  $R^{33}(CH_2)_kC(O)-$  or  $H-C(O)-$ ; where  $k$  is an integer from 1 to 4, or is 1 or 2;
- 20  $R^{30}$  is alkyl, aryl, heteroaryl, aralkyl or heteroaralkyl;  $R^{31}$  and  $R^{32}$  are each independently hydrogen, alkyl, aryl, heteroaryl, aralkyl, or heteroaralkyl; and  $R^{33}$  is hydrogen, hydroxy, alkyl, alkenyl, alkynyl, alkoxy, aryl, aryloxy, heteroaryl, heteroaryloxy, aralkyl, aralkoxy, heteroaralkyl or heteroaralkoxy.

- 25 In certain embodiments herein,  $X^n$  is hydrogen, acetyl, hydroxyacetyl, 2,3-dihydroxypropionyl, 2,3,4-trihydroxybutanoyl, PEG(1), PEG(2), PEG(4), PEG(6), PEG(14), PEG(15), PEG(16), PEG(17), PEG(18) or PEG(19). In other embodiments herein,  $X^n$  is hydrogen, acetyl, hydroxyacetyl, succinyl, quinyll, gallyl, 4-imidazolylacetyl, cotininyll, 3-phosphonylpropionyl, gulonyl, 4-phosphonylbutyryl, glutaryl,
- 30 ethoxysquaryl or PEG(2). In further embodiments,  $X^n$  is hydrogen, acetyl,  $-C(O)NH_2$ ,  $HOCH_2CH_2C(O)-$ , diaminopropanoyl, or  $NH_2-(CH_2)_5-C(O)-$ . In another embodiment,  $X^n$  is hydrogen, acetyl, succinyl, glutaryl, PEG(2) or malonyl. In another embodiment,  $X^n$  is hydrogen, acetyl, succinyl,

-95-

glutaryl, PEG(2), malonyl, methoxycarbonyl, phenylsulfonyl, 3-methoxypropanoyl, ethoxycarbonyl, isobutoxycarbonyl, benzyloxycarbonyl, tert-butoxycarbonyl, 4-oxopentanoyl, 2-(2-methoxyethoxy)ethoxy)acetyl, 3,4-methylenedioxyphenylacetyl, 2-pyridylacetyl, phenoxyacetyl, phenylacetyl, methoxyacetyl, 2-methoxyethoxycarbonyl, 2-methoxyethoxyacetyl, 3-phenyl-2-hydroxypropanoyl, pent-4-ynoyl, 1-naphthylacetyl, hydroxyacetyl, 3-methoxycarbonylpropanoyl or formyl.

In certain embodiments herein, the N-terminal cap ( $X^n$ ) is acetyl, glutaryl, or related acyl, sulfonyl or carbamoyl derivatives. Capping groups include, but are not limited to, a simple N-acetyl residue through larger fragments that impact the overall physicochemical properties of the conjugate. Appropriate choice of the capping group allows delivery of either relatively hydrophilic or hydrophobic molecules to a target site. In one embodiment,  $X^n$  is acetyl.

## 2) $X^c$ (the C-terminal Cap)

In embodiments herein where the therapeutic agent is conjugated to the N-terminus of the peptidic substrate (*i.e.*, where the conjugate has formula II), the C-terminus of the peptidic substrate is a carboxylic acid or a carboxamide derivative. Appropriate choice of the capping group allows delivery of either relatively hydrophilic or hydrophobic molecules to a target site.

In one embodiment,  $X^c$ , together with the carbonyl group to which it is attached, forms a carboxamide derivative of formula  $-C(O)NR^dR^e$ , where  $R^d$  and  $R^e$  are selected from (i) or (ii) as follows:

(i)  $R^d$  and  $R^e$  are each independently hydrogen,  $C_1$ - $C_6$ -alkyl,  $-C_1$ - $C_6$ -alkyl-OH,  $-C_1$ - $C_6$ -alkyl-di-OH,  $-C_1$ - $C_6$ -alkyl-tri-OH and



-96-

provided that at least one of  $R^d$  and  $R^e$  are not hydrogen or  $C_1$ - $C_6$ -alkyl; or

(ii)  $R^d$  and  $R^e$  together form a  $-CH_2CH_2OCH_2CH_2-$  diradical;

b is zero or an integer between 1 and 100; and

c is 0 or 1, provided that if b is zero, c is 1.

5 In one embodiment,  $R^d$  is hydrogen and  $R^e$  is 2-hydroxyethyl.

## 2. Linkers

The conjugates optionally contain a linker (*i.e.*, L,  $L^1$  or  $L^2$  of formulae I, II and III) that covalently binds the peptidic substrate to the therapeutic agent. The linkers are any that result in a conjugate in which  
10 the peptidic portion is a substrate for a cell surface protease and the therapeutic agent is substantially inactive when in the conjugate and is released in active form or in a form subsequently activated by the cell, tissue or environment of the targeted tissue.

For example, the linker can include of carbohydrate, peptide,  
15 diamine, arylamine, and/or hydrocarbon core structures. Linkers are desirably synthetically accessible, provide shelf-stable products, and do not possess any intrinsic biological activity that interferes with the conjugates activity. They can add desirable properties such as increasing solubility or serving to aid in trafficking the cleaved therapeutic agent in  
20 the cell. In certain embodiments, some linkers will be enzymatically cleaved *in vitro* and *in vivo*, and fragment to release active therapeutic agent or activatable therapeutic agent. In embodiments where the therapeutic agent is doxorubicin, the linker is, for example, a sugar and/or a peptide, such the aminosugar daunosamine.

25 In one embodiment, linkers for use herein include, but are not limited to, a biscarbonyl alkyl diradical whereby an amine moiety on the therapeutic agent is connected with the linker unit to form an amide bond and the amino terminus of the peptidic substrate is connected with the other end of the linker unit also forming an amide bond. Conversely, a

-97-

diaminoalkyl diradical linker unit, whereby a carbonyl moiety on the cytotoxic agent is covalently attached to one of the amines of the linker unit while the other amine of the linker unit is covalently attached to the C-terminus of the peptidic substrate, also can be useful. Other such linker units which are stable to the physiological environment when not in the presence of a cell surface protease, but are cleavable upon the cleavage of the cell surface protease proteolytic cleavage site, are intended for use herein. Furthermore, linker units can be utilized that, upon cleavage of the cell surface protease proteolytic cleavage site, remain attached to the therapeutic agent but do not significantly decrease the therapeutic activity of such a post-cleavage therapeutic agent derivative when compared with an unmodified therapeutic agent.

In other embodiments, the linker is a diamine containing a cyclic alkyl moiety and, in certain embodiments, the diamine contains a bicycloalkylene moiety. Examples of such diamine linkers include, but are not limited to, 1,4-bis(aminomethyl)cyclohexane, 1,4-bis(aminomethyl)-cycloheptane, 1,3-bis(aminomethyl)cyclopentane, 1-amino-4-(aminomethyl)cyclohexane, 1,4-diaminocyclohexane and 1,4-bis(aminomethyl)-bicyclo[2.2.2]octane.

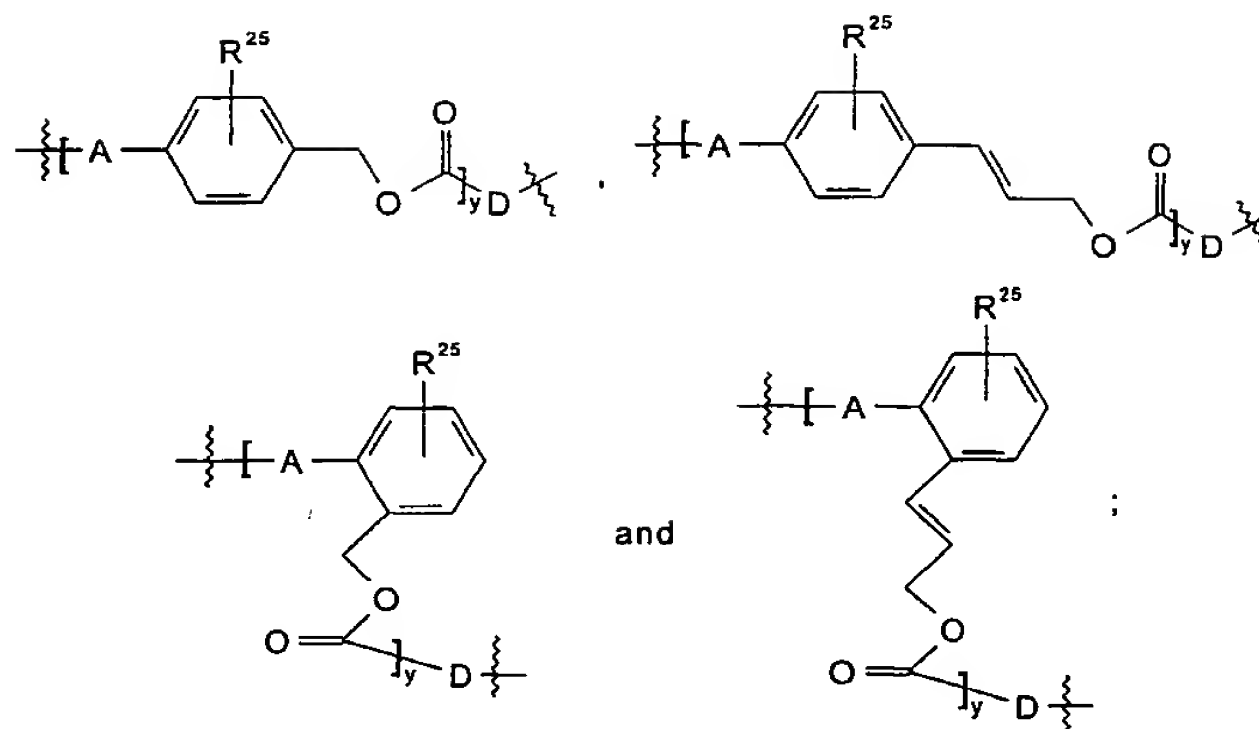
Other linkers include 1, $\omega$ -diaminoalkanes, including, but not limited to, 1,3-diaminopropane, and 1, $\omega$ -dicarbonylalkanes, including, but not limited to, oxalic, malonic, succinic, glutaric, adipic and pivalic acids.

Further linkers for use in the conjugates provided herein include self-eliminating linkers such as those of the following formulae:

25

30

-98-



- where A is NH or O; D is N(H or alkyl) or O; R<sup>25</sup> is H, alkyl, cycloalkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, such as 1 to 3, substituents selected from halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2 triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy, haloalkyl and polyhaloalkyl, such as, for example, halo lower alkyl, especially trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that optionally is substituted with 1 or more, such as, for example, 1 to 3, substituents selected from, for example, halo, halo alkyl and alkyl, heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxy, aminoimino, alkoxycarbonylamino, aryloxy, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, arylaminocarbonyl, diarylaminocarbonyl, aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkyl-aminoalkyl, dialkylaminoalkyl, arylaminoalkyl, amino, alkylamino, dialkyl-amino, arylamino, alkylaryl amino, alkylcarbonylamino, arylcarbonylamino, azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyno,

-99-

isothiocyano, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl, alkylaminosulfonyl, dialkylaminosulfonyl and arylamino-sulfonyl.; and y is an integer from 1 to 3.

### 3. Therapeutic agents

5 The conjugates are intended for modifying a variety of biological responses. Accordingly, the therapeutic agents are any agents, including proteins and polypeptides, small molecules and other molecules that possess or potentiate a desired biological activity. Such molecules include cytotoxic agents, such as, but are not limited to, a toxin such as  
10 abrin, ricin A, pseudomonas exotoxin, shiga toxin, diphtheria toxin and other such toxins and toxic portions and/or subunits or chains thereof; proteins such as, but not limited to, tumor necrosis factor,  $\alpha$ -interferon,  $\gamma$ -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for  
15 example, lymphokines, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), erythropoietin (EPO), pro-coagulants such as tissue factor and tissue factor variants, pro-apoptotic agents such as FAS-ligand, fibroblast growth factors (FGF),  
20 nerve growth factor and other growth factors. Each must be in a form that can enter a cell or otherwise exert a therapeutic effect when in the vicinity thereof.

Thus, therapeutic agents, include, but are not limited to, anti-tumor, anti-angiogenic, pro-apoptotic, anti-cancer and anti-mitotic agents.  
25 These are conjugated, optionally via a linker, to a substrate, such as peptidic substrate, which is a substrate for the protease.

Among the therapeutic agents are cytotoxic agents that include, in general, but are not limited to, alkylating agents, toxins, antiproliferative agents and tubulin binding agents. Classes of cytotoxic agents for use



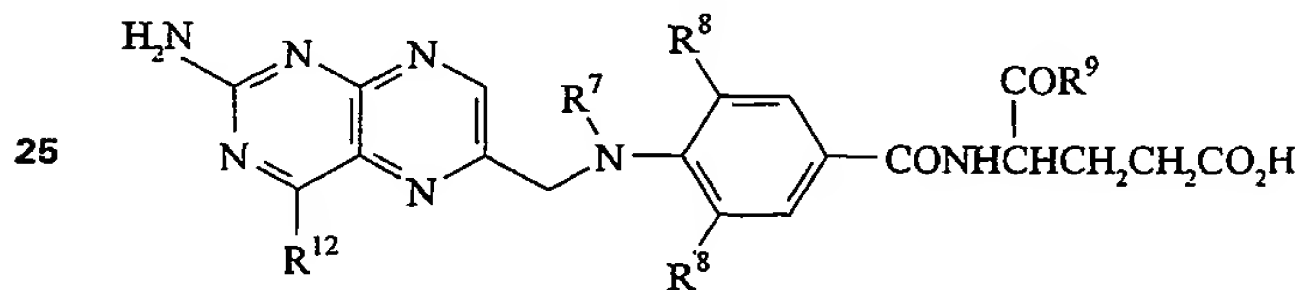
-100-

herein include, for example, the anthracycline family of drugs, the vinca drugs, the mitomycins, the bleomycins, the cytotoxic nucleosides, the pteridine family of drugs, diylenes, the maytansinoids, the epothilones, the taxanes and the podophyllotoxins.

5 Exemplary members of those classes include, for example, doxorubicin, carminomycin, daunorubicin, aminopterin, methotrexate, methopterin, dichloro-methotrexate, mitomycin C, porfiromycin, 5-fluorouracil, 6-mercaptopurine, cytosine arabinoside, podophyllotoxin, or podophyllotoxin derivatives such as etoposide or etoposide phosphate,  
 10 melphalan, vinblastine, vincristine, leurosine, vindesine, leurosine, maytansinol, epothilone A or B, taxotere, taxol and the like. Other such therapeutic agents include estramustine, cisplatin, combretastatin and analogs, and cyclophosphamide. One skilled in the art can make chemical modifications to the desired therapeutic agent in order to make  
 15 reactions of that compound more convenient for purposes of preparing the conjugates.

Particular therapeutic agents include the following drugs. One skilled in the art understands that these structural formulae are exemplary only and that such compounds or derivatives or analogs thereof have  
 20 acquired in the art different generic or trivial names.

a. The methotrexate group of formula (1):



30 in which

$R^{12}$  is amino or hydroxy;

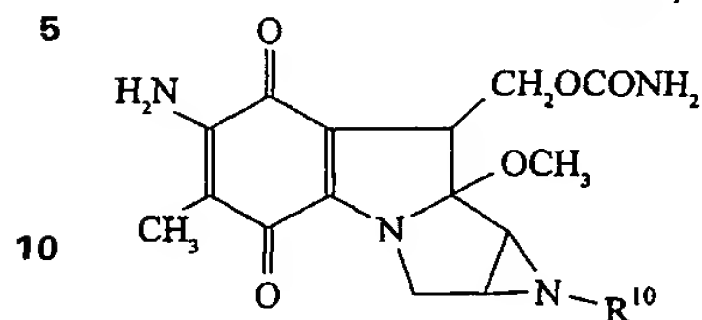
$R^7$  is hydrogen or methyl;

-101-

$R^8$  is hydrogen, fluoro, chloro, bromo or iodo;

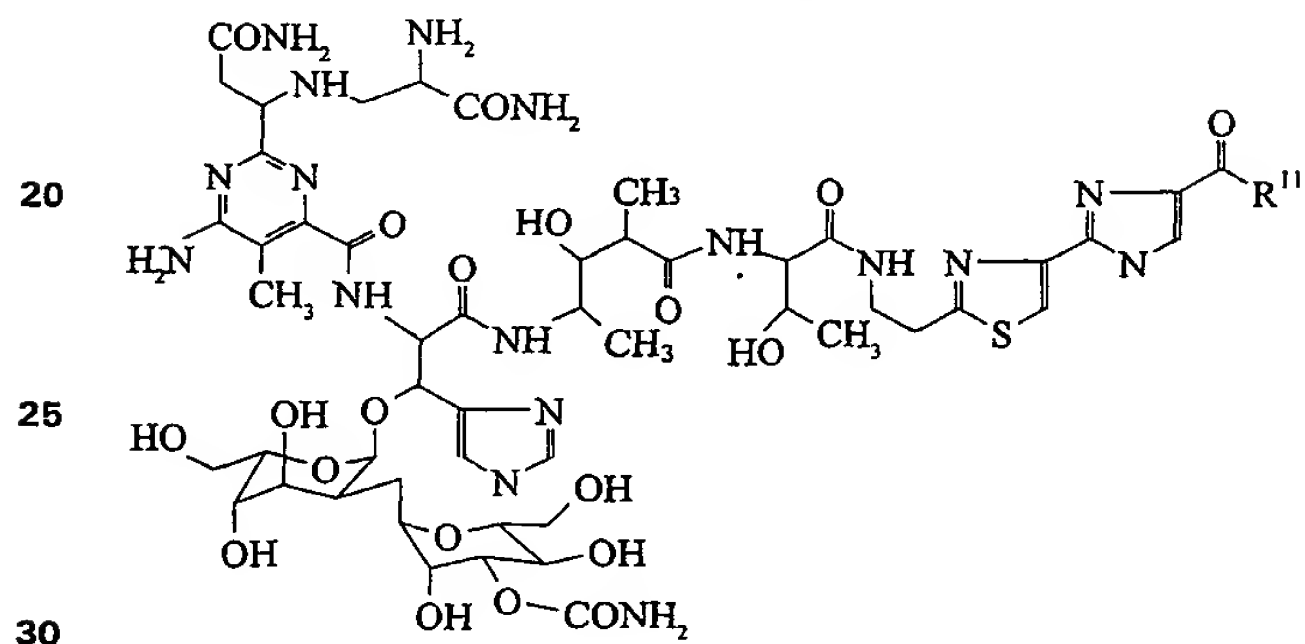
$R^9$  is hydroxy or a moiety which completes a salt of the carboxylic acid.

**b. The mitomycin group of formula (2):**



in which  $R^{10}$  is hydrogen or methyl.

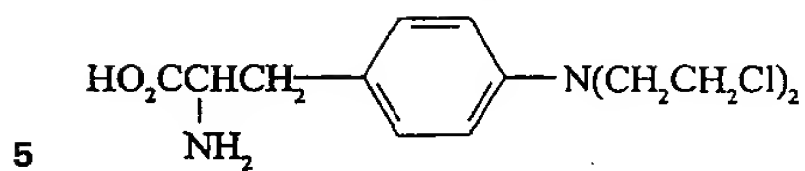
**c. The bleomycin group of formula (3):**



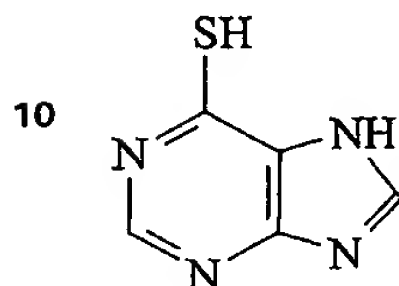
in which  $R^{11}$  is hydroxy, amino, C<sub>1</sub>-C<sub>3</sub> alkylamino, di(C<sub>1</sub>-C<sub>3</sub> alkyl)amino, C<sub>4</sub>-C<sub>6</sub> polymethylene amino, -NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH-C(NH)NH<sub>2</sub> or -NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>S<sup>+</sup>(CH<sub>3</sub>)<sub>2</sub>.

-102-

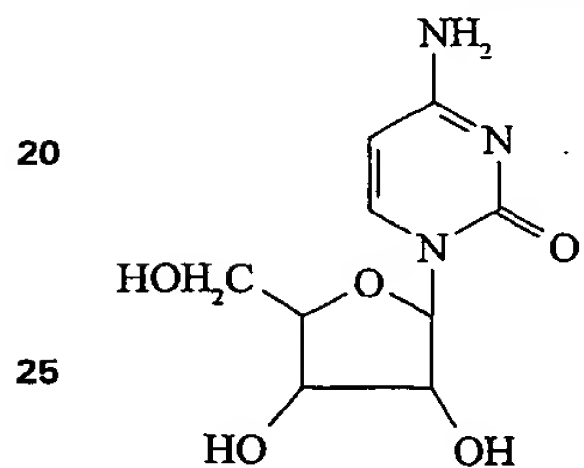
d. Melphalan of formula (4):



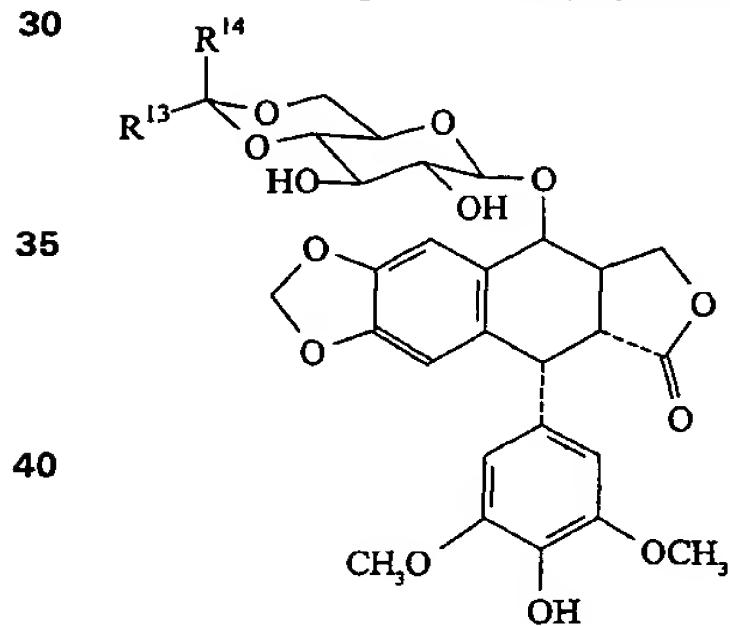
e. Mercaptopurine of formula (5):



f. Cytosine arabinoside of formula (6):



g. Podophyllotoxins of formula (7):



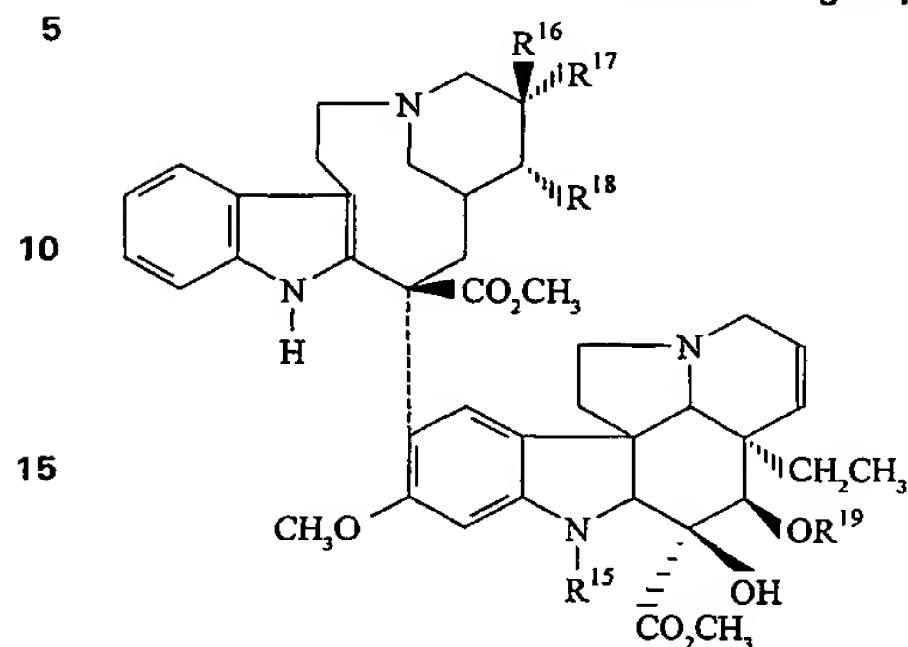
-103-

in which

$R^{13}$  is hydrogen or methyl; and

$R^{14}$  is methyl or thienyl or a phosphate salt thereof.

**h. The vinca alkaloid group of drugs of formula (8):**



in which

when  $R^{17}$  and  $R^{18}$  are taken singly,  $R^{15}$  is H,  $CH_3$  or CHO; and

$R^{18}$  is H, and one of  $R^{16}$  and  $R^{17}$  is ethyl and the other is H or OH;

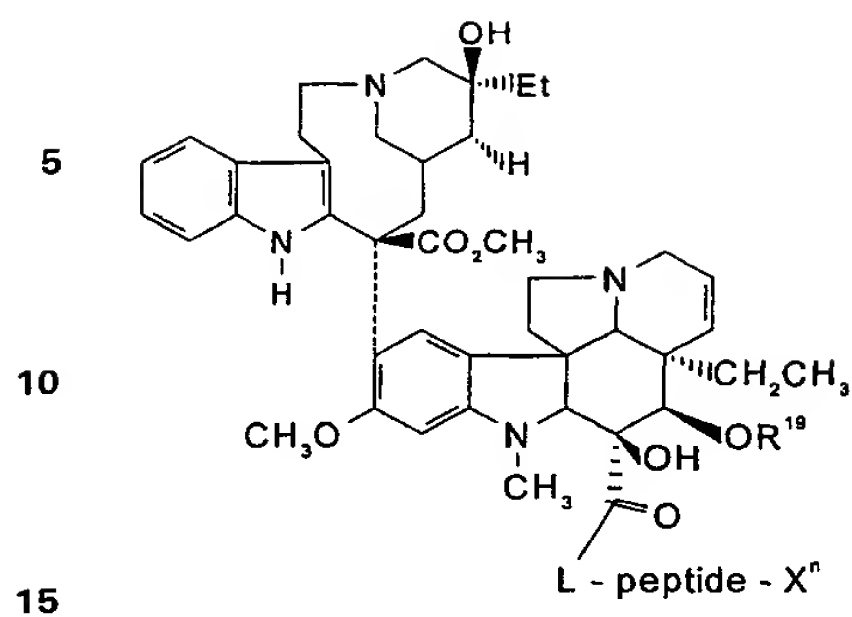
25 when  $R^{17}$  and  $R^{18}$  are taken together with the carbons to which they are attached, they form an oxirane ring in which case  $R^{16}$  is ethyl; and

$R^{19}$  is hydrogen,  $(C_1-C_3 \text{ alkyl})-CO$ , or chlorosubstituted  $(C_1-C_3 \text{ alkyl})-CO$ .

30 The conjugates provided herein where the therapeutic agent is the vinca alkaloid vinblastine include those of formula:

35

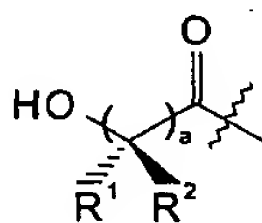
-104-



where the peptidic substrate is as described above for formulae I and II; L is a linker such as -NH-(CH<sub>2</sub>)<sub>u</sub>-T-(CH<sub>2</sub>)<sub>v</sub>-NH-; X<sup>n</sup> is

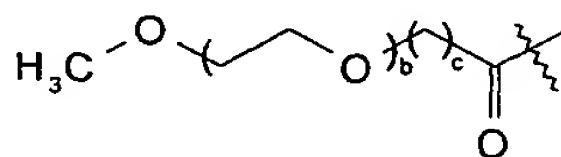
- 20
- a) hydrogen,
- b) -(C=O)R<sup>1a</sup>,
- c)

25



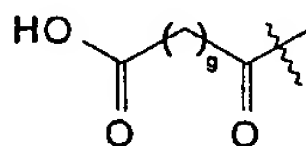
30

d)



35

e)



40

- f) ethoxysquarate; and

-105-

g) cotininyl;

$R^1$  and  $R^2$  are independently hydrogen, OH,  $C_1$ - $C_6$  alkyl,  $C_1$ - $C_6$  alkoxy,  $C_1$ - $C_6$  aralkyl and aryl;

$R^{1a}$  is  $C_1$ - $C_6$ -alkyl, hydroxylated  $C_3$ - $C_8$ -cycloalkyl, polyhydroxylated  $C_3$ - $C_8$ -cycloalkyl, hydroxylated aryl, polyhydroxylated aryl or aryl,

$R^{19}$  is hydrogen, ( $C_1$ - $C_3$  alkyl)-CO, or chlorosubstituted ( $C_1$ - $C_3$  alkyl)-CO;

T is selected from cyclopentyl, cyclohexyl, cycloheptyl or bicyclo[2.2.2]octanyl;

10 a is 1, 2, 3 or 4;

b is zero or an integer between 1 and 100;

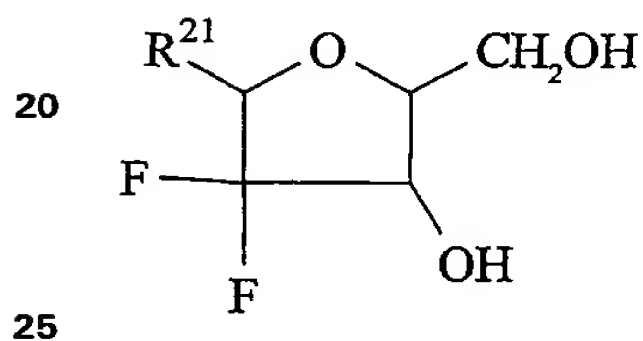
c is 0 or 1, provided that if b is zero, c is 1;

g is 1, 2 or 3;

u is 0, 1, 2 or 3;

15 or a pharmaceutically acceptable derivative thereof.

i. Difluoronucleosides of formula (9):

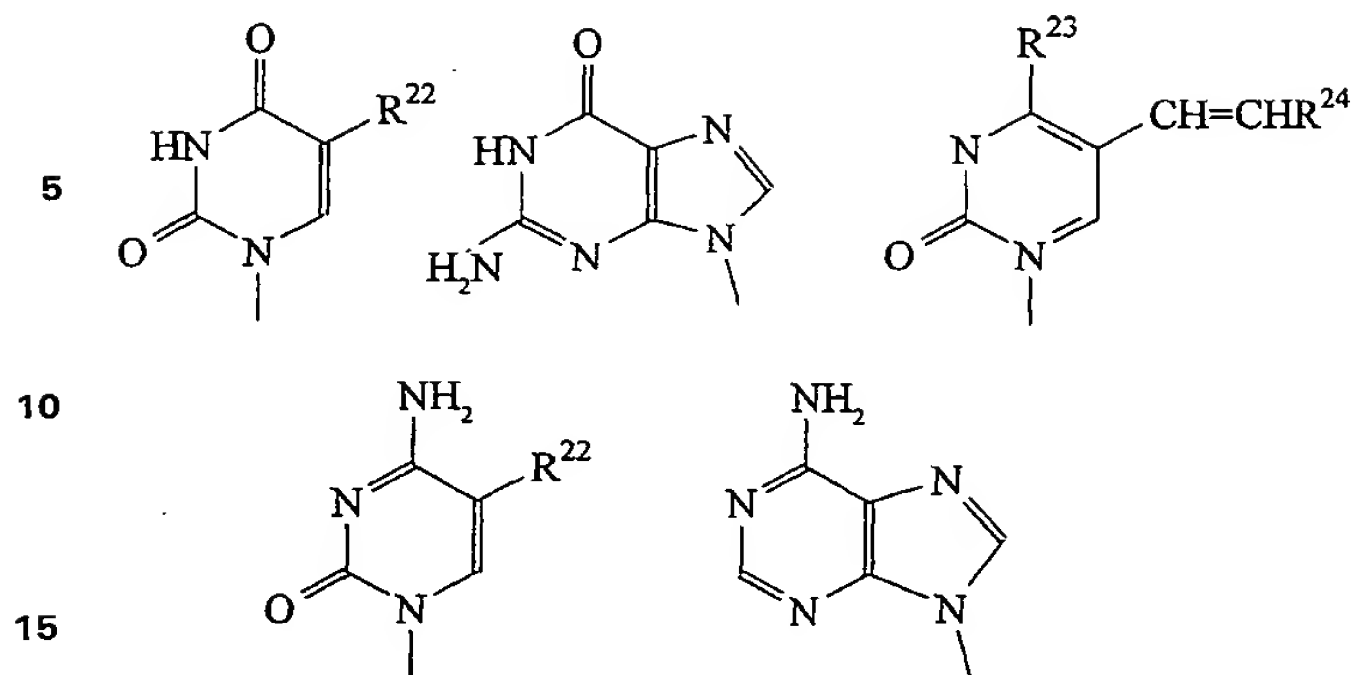


in which  $R^{21}$  is a base of one of the formulae:

30

35

-106-

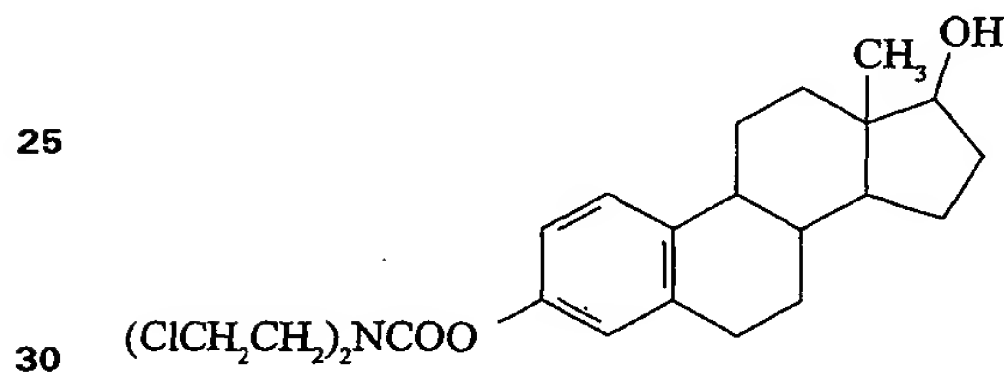


$R^{22}$  is hydrogen, methyl, bromo, fluoro, chloro or iodo;

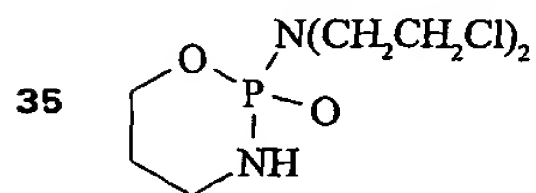
$R^{23}$  is -OH or -NH<sub>2</sub>;

20  $R^{24}$  is hydrogen, bromo, chloro or iodo.

j. Estramustine (10):

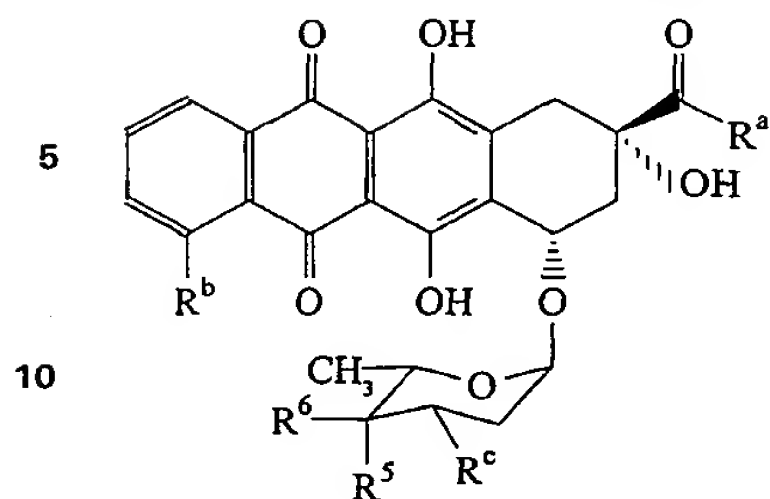


k. Cyclophosphamide (11):



-107-

## I. Anthracycline antibiotics of formula (12):



15 in which

$R^a$  is  $-\text{CH}_3$ ,  $-\text{CH}_2\text{OH}$ ,  $-\text{CH}_2\text{OCO}(\text{CH}_2)_3\text{CH}_3$ , or  
 $-\text{CH}_2\text{OCOCH}(\text{OC}_2\text{H}_5)_2$ ;

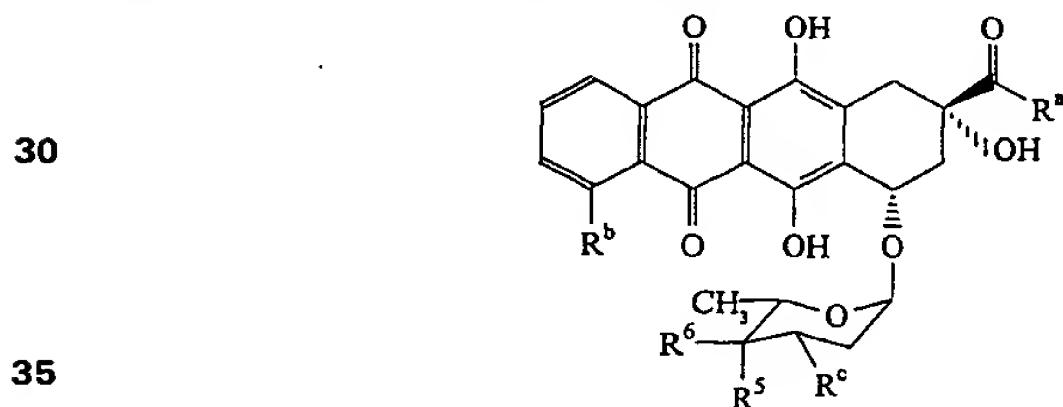
$R^b$  is  $-\text{OCH}_3$ ,  $-\text{OH}$  or  $-\text{H}$ ;

$R^c$  is  $-\text{NH}_2$ ,  $-\text{NHCOCF}_3$ , 4-morpholinyl, 3-cyano-4-morpholinyl,  
 20 1-piperidinyl, 4-methoxy-1-piperidinyl, benzylamine, dibenzylamine,  
 cyanomethylamine, or 1-cyano-2-methoxyethyl amine;

$R^5$  is  $-\text{OH}$ ,  $-\text{OTHP}$  or  $-\text{H}$ ; and

$R^6$  is  $-\text{OH}$  or  $-\text{H}$  provided that  $R^6$  is not  $-\text{OH}$  when  $R^5$  is  $-\text{OH}$  or  
 $-\text{OTHP}$ .

25 Table 2, which follows, provides a number of anthracycline drugs  
 and their generic or trivial names:



Compound	$R^a$	$R^b$	$R^c$	$R^5$	$R^6$
----------	-------	-------	-------	-------	-------



-108-

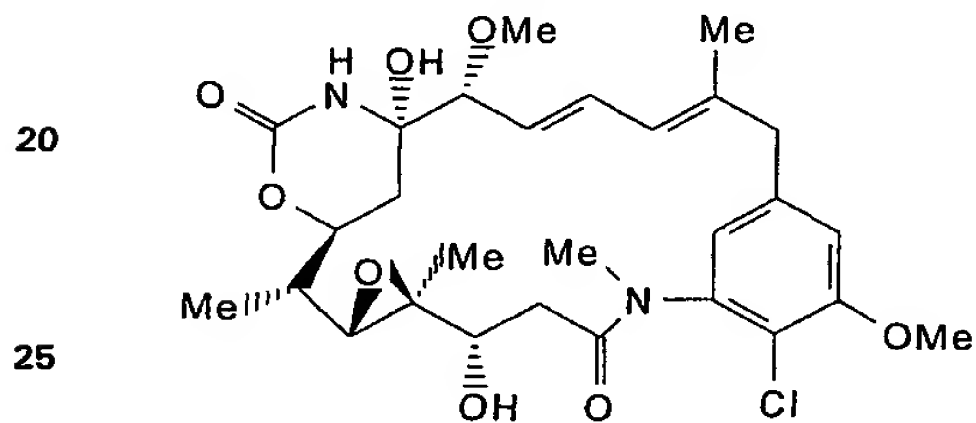
daunorubicin <sup>a</sup>	CH <sub>3</sub>	OCH <sub>3</sub>	NH <sub>2</sub>	OH	H
doxorubicin <sup>b</sup>	CH <sub>2</sub> OH	OCH <sub>3</sub>	NH <sub>2</sub>	OH	H
detorubicin	CH <sub>2</sub> OCOCH(OC <sub>2</sub> H <sub>5</sub> ) <sub>2</sub>	OCH <sub>3</sub>	NH <sub>2</sub>	OH	H
carminomycin	CH <sub>3</sub>	OH	NH <sub>2</sub>	OH	H
5 idarubicin	CH <sub>3</sub>	H	NH <sub>2</sub>	OH	H
epirubicin	CH <sub>2</sub> OH	OCH <sub>3</sub>	NH <sub>2</sub>	OH	OH
esorubicin	CH <sub>2</sub> OH	OCH <sub>3</sub>	NH <sub>2</sub>	H	H
THP	CH <sub>2</sub> OH	OCH <sub>3</sub>	NH <sub>2</sub>	OTHP	H
10 AD-32	CH <sub>2</sub> OCO(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	OCH <sub>3</sub>	NHCOCF <sub>3</sub>	OH	H

<sup>a</sup> daunorubicin is an alternative name for daunomycin

<sup>b</sup> doxorubicin is an alternative name for adriamycin.

In one embodiment, when the therapeutic agent is doxorubicin, it is conjugated to the peptidic substrate via the amino group of the aminoglycoside moiety of doxorubicin.

m. Maytansinol

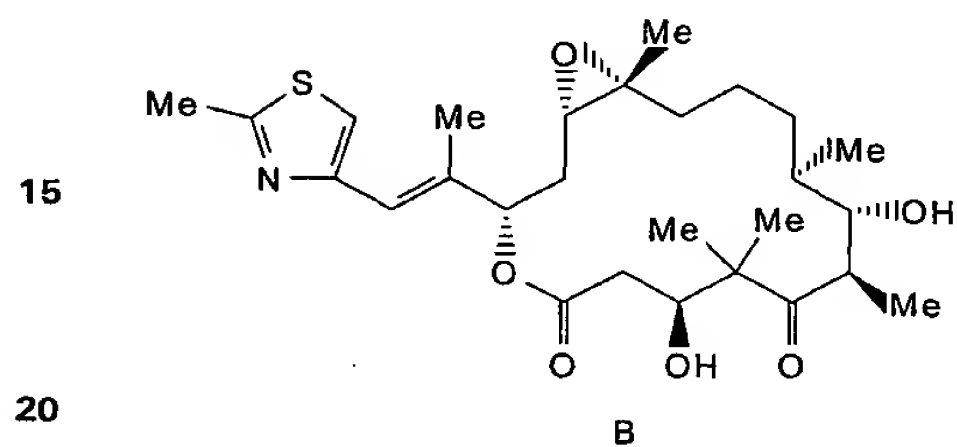
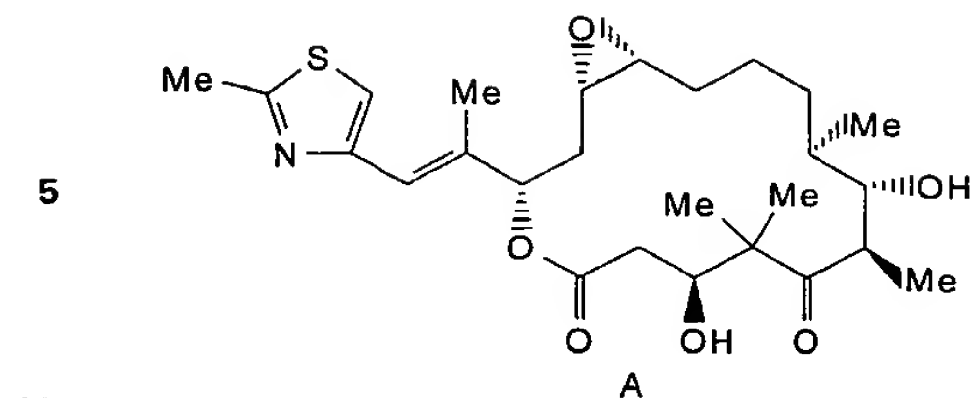


n. Epothilone A or B

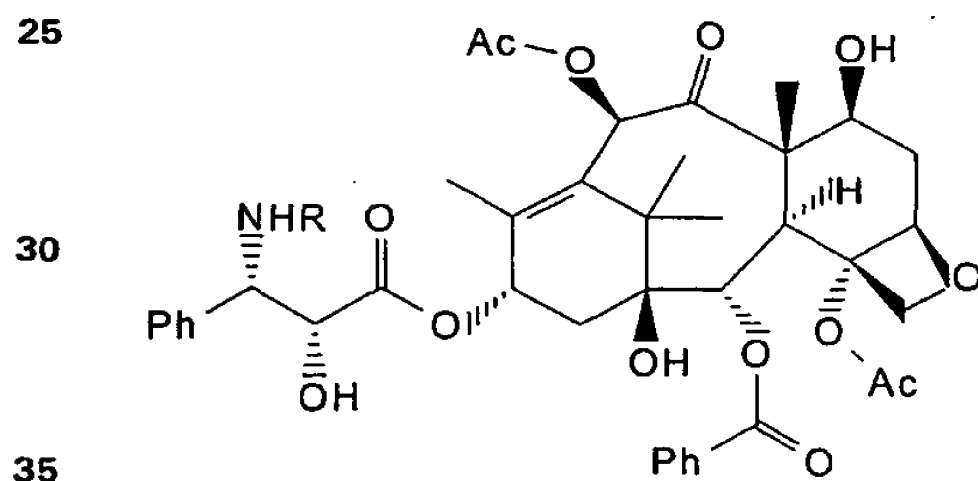
35

40

-109-



o. Taxols



where R is PhC(O) or t-BuOC(O).

In one embodiment, when the therapeutic agent is taxol (R = C(O)Ph), the peptidic substrate is conjugated to the secondary hydroxyl  
 40 group of the cyclohexane moiety of taxol.

-110-

**p. Ribosome-inactivating proteins**

Ribosome-inactivating proteins (RIPs), which include ricin, abrin and saporin, are plant proteins that catalytically inactivate eukaryotic ribosomes. RIPS inactivate ribosomes by interfering with the protein  
5 elongation step of protein synthesis. For example, the RIP saporin (hereinafter also referred to as SAP) has been shown to enzymatically inactivate 60S ribosomes by cleavage of the n-glycosidic bond of the adenine at position 4324 in the rat 28S ribosomal RNA (rRNA). Some RIPs, such as the toxins abrin and ricin, contain two constituent chains:  
10 a cell-binding chain that mediates binding to cell surface receptors and internalization of the molecule; and an enzymatically active chain responsible for protein synthesis inhibitory activity. Such RIPs are type II RIPs. Other RIPs, such as the saporins, are single chains and are designated type I RIPs. Because such RIPs lack a cell-binding chain, they  
15 are less toxic to whole cells than the RIPs that have two chains. Two chain RIPs are generally used for conjugation herein, unless a single chain is further conjugated to an agent, such as a growth factor that mediates binding and internalization.

Several structurally related RIP's have been isolated from seeds and  
20 leaves of the plant *Saponaria officinalis* (soapwort). Among these, SAP-6 is the most active and abundant, representing 7% of total seed proteins. Saporin is very stable, has a high isoelectric point, does not contain carbohydrates, and is resistant to denaturing agents, such as sodium dodecyl sulfate (SDS), and a variety of proteases. The amino acid  
25 sequences of several saporin-6 isoforms from seeds are known and there appear to be families of saporin RIPs differing in a few amino acid residues. Because saporin is a type I RIP, it does not possess a cell-binding chain. Consequently, its toxicity to whole cells is much lower than the other toxins, such as ricin and abrin. When internalized by

-111-

eukaryotic cells, however, its cytotoxicity is 100- to 1000-fold more potent than ricin A chain.

#### 4. Exemplary Conjugates

The conjugates provided herein, are prepared by identifying suitable  
5 peptidic substrates for the targeted cell surface protease, or a soluble, shed or released form thereof, and forming a conjugate of the peptidic substrate(s) with a therapeutic agent(s). Exemplary conjugates, containing peptidic substrates designed, for example, for cleavage by MTSP1, endotheliase 1 and urokinase, are described. It is understood  
10 that upon identification of a cell surface protease, including cell-associated and cell-localized proteases, or a soluble, shed or released form thereof, in or associated with a cell involved in a disease or other conditions of interest, or with a cell present in the vicinity of a cell or tissue involved in or associated with a disease or other condition of  
15 interest, suitable peptidic substrates therefor can be empirically designed and then conjugated to therapeutic agents as exemplified herein.

In certain embodiments, the conjugates for use in the compositions and methods provided herein include:

- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO:  
20 46);  
Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 47);  
Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 48);  
25 Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 49);  
Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 50);  
Ac-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 51);  
Ac-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 52);

-112-

Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 53);  
 Ac-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 54);  
 Ac-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 55);  
 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 56);  
 5 Ac-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 57);  
 Ac-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 58);  
 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 59);  
 Ac-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 60); and  
 Ac-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 61).

10 In further embodiments herein, the conjugates are Ac-Leu-Arg-Ala-  
 Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 62); Ac-Leu-Arg-  
 Ala-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 63); Ac-Leu-  
 Arg-Ser-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 64); and  
 Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO:  
 15 65).

In other embodiments herein, the conjugates are  
 Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 66);  
 Ac-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 67);  
 Ac-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 68);  
 20 Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 69);  
 Ac-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 70);  
 Ac-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 71);  
 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 72);  
 Ac-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 73);  
 25 Ac-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 74);  
 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 75);  
 Ac-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 76); and  
 Ac-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 77).

-113-

In other embodiments, the conjugates for use herein include the following:

- pyroGlu-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 78);
- CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 79);
- 5 N-p-tosyl-Gly-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 80);
- Benzoyl-Val-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 81);
- CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 82);
- N- $\alpha$ -Z-D-Arg-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 83) (Z = benzyloxycarbonyl);
- 10 pyroGlu-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 84);
- H-D-Ile-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 85);
- Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 86) (Cbo = carbobenzoxy);
- H-D-Pro-Phe-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 87);
- 15 H-D-Val-Leu-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 88);
- Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 89) (Bz = benzoyl);
- Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 90);
- Benzoyl-Pro-Phe-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 91);
- 20 H-D-Phe-Pip-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 92);
- H-D-Val-Leu-Lys-Ala-Ala-(therapeutic agent) (SEQ ID NO: 93);
- H-D-Nle-HHT-Lys-Ala-Ala-(therapeutic agent) (SEQ ID NO: 94);
- Pyr-Arg-Thr-Lys-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 95);
- H-Arg-Gln-Arg-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 96);
- 25 Boc-Gln-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 97);
- Z-Arg-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 98);
- H-D-HHT-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 99);
- H-D-CHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 100);
- MeSO<sub>2</sub>-dPhe-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 101);

-114-

$\delta$ -Z-D-Lys-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 102); and  
CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 103).

In another embodiment, the conjugates for use in the compositions and methods provided herein include:

- 5 Ac-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 104);  
Ac-Arg-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 105);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 106);  
Ac-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 107);  
10 Ac-Arg-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 108);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Gly-Gly-(therapeutic agent) (SEQ ID NO: 109);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 110);  
Ac-Arg-Arg-Gln-Ser-Arg-Ile-(therapeutic agent) (SEQ ID NO: 111); and  
15 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-Ile-(therapeutic agent) (SEQ ID NO: 112).

In certain embodiments, the conjugates for use in the compositions and methods provided herein include:

- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 113);  
20 Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 114);  
Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 115);  
Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 116);  
25 Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 117);  
Ac-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 118);  
Ac-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 119);

-115-

Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 120);

Ac-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 121);

Ac-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 122);

5 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 123);

Ac-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 124);

Ac-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 125);

10 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 126);

Ac-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 127);

and

Ac-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 128).

15 In further embodiments herein, the conjugates are Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 129); Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 130); Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 131); and Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 132).

20 In other embodiments herein, the conjugates are

Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 133);

Ac-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 134);

Ac-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 135);

25 Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 136);

Ac-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 137);

Ac-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 138);



-116-

Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 139);

Ac-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 140);

Ac-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 141);

5 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 142);

Ac-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 143);

and

Ac-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 144).

10 In other embodiments, the conjugates for use herein include the following:

pyroGlu-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 145);

CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 146);

N-p-tosyl-Gly-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 147);

15 Benzoyl-Val-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 148);

CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 149);

N- $\alpha$ -Z-D-Arg-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 150) (Z = benzyloxycarbonyl);

pyroGlu-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 151);

20 H-D-Ile-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 152);

Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 153) (Cbo = carbobenzoxy);

H-D-Pro-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 154);

H-D-Val-Leu-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 155);

25 Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 156) (Bz = benzoyl);

Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 157);

Benzoyl-Pro-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 158);

H-D-Phe-Pip-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 159);

-117-

- H-D-Val-Leu-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 160);  
 H-D-Nle-HHT-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 161);  
 Pyr-Arg-Thr-Lys-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 162);  
 H-Arg-Gln-Arg-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 163);  
 5 Boc-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 164);  
 Z-Arg-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 165);  
 H-D-HHT-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 166);  
 H-D-CHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 167);  
 MeSO<sub>2</sub>-dPhe-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 168);  
 10  $\delta$ -Z-D-Lys-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 169); and  
 CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 170).

In another embodiment, the conjugates for use in the compositions and methods provided herein include:

- Ac-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 171);  
 15 Ac-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 172);  
 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 173);  
 Ac-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 174);  
 Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 175);  
 20 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 176);  
 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 177);  
 Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 178); and  
 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 179).  
 25

In other embodiments, the conjugates provided herein include:

- Ac-Arg-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 180);  
 Ac-Arg-Gln-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 181);  
 Ac-Arg-Gln-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 182);

-118-

- Ac-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 183);  
Ac-Arg-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 184);  
Ac-Arg-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 185);  
Ac-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 186);  
5 Ac-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 187);  
Ac-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 188);  
Ac-Gln-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 189); and  
Ac-Gln-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 190).

- In further embodiments, the conjugates for use in the compositions  
10 and methods provided herein include:  
Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
191);  
Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
192);  
15 Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
193);  
Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
194);  
Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
20 195);  
Ac-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
196);  
Ac-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 197);  
Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
25 198);  
Ac-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
199);  
Ac-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 200);

-119-

Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 201);

Ac-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 202);

- 5 Ac-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 203);  
Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 204);

Ac-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 205); and

- 10 Ac-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 206).

- In further embodiments herein, the conjugates are Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 207); Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 208); Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 209); and Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 210).

In other embodiments herein, the conjugates are

Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 211);

- 20 Ac-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 212);

Ac-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 213);

Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 214);

- 25 Ac-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 215);

Ac-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 216);

Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 217);

-120-

Ac-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 218);

Ac-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 219);

Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 220);

Ac-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 221); and

Ac-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 222).

In other embodiments, the conjugates for use herein include the following:

pyroGlu-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 223);

CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 224);

N-p-tosyl-Gly-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 225);

Benzoyl-Val-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 226);

CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 227);

N- $\alpha$ -Z-D-Arg-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 228) (Z = benzyloxycarbonyl);

pyroGlu-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 229);

H-D-Ile-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 230);

Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 231) (Cbo = carbobenzoxy);

H-D-Pro-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 232);

H-D-Val-Leu-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 233);

Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 234) (Bz = benzoyl);

Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 235);

-121-

- Benzoyl-Pro-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 236);  
 H-D-Phe-Pip-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 237);  
 H-D-Val-Leu-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 238);  
 H-D-Nle-HHT-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 239);  
 5 Pyr-Arg-Thr-Lys-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 240);  
 H-Arg-Gln-Arg-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 241);  
 Boc-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 242);  
 Z-Arg-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 243);  
 H-D-HHT-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 244);  
 10 H-D-CHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 245);  
 MeSO<sub>2</sub>-dPhe-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 246);  
 δ-Z-D-Lys-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 247); and  
 CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
 248).  
 15 In another embodiment, the conjugates for use in the compositions  
 and methods provided herein include:  
 Ac-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 249);  
 Ac-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
 250);  
 20 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
 251);  
 Ac-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 252);  
 Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 253);  
 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
 25 254);  
 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO:  
 255);  
 Ac-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 256);  
 and

-122-

Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 257).

In other embodiments, the conjugates provided herein include:

- Ac-Arg-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 258);
- 5 Ac-Arg-Gln-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 259);
- Ac-Arg-Gln-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 260);
- Ac-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 261);
- Ac-Arg-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 262);
- Ac-Arg-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 263);
- 10 Ac-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 264);
- Ac-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 265);
- Ac-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 266);
- Ac-Gln-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 267); and
- Ac-Gln-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 268).

- 15 In another embodiment, the conjugates provided herein include:
- Ac-Gly-dSer-Ala-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 569);
- Ac-Arg-Gly-dSer-Ala-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 570);
- Ac-Gly-Ser-Gly-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 571);
- Ac-Arg-Gly-Ser-Gly-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 572);
- 20 Ac-Leu-Arg-Gly-Ser-Gly-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 573);
- Ac-Leu-Arg-Gly-dSer-Ala-Arg-Ser-Ala-(therapeutic agent) (SEQ ID NO: 574);
- Ac-Cys(Me)-Pro-Gly-Arg-Val-Val-(therapeutic agent) (SEQ ID NO: 575);
- 25 Ac-Arg-Cys(Me)-Pro-Gly-Arg-Val-Val-(therapeutic agent) (SEQ ID NO: 577);
- Ac-Arg-Arg-Cys(Me)-Pro-Gly-Arg-Val-Val-(therapeutic agent) (SEQ ID NO: 578);
- Ac-Val-Ser-Ala-Arg-Met-Ala-(therapeutic agent) (SEQ ID NO: 579);

-123-

- Ac-Ile-Val-Ser-Ala-Arg-Met-Ala-(therapeutic agent) (SEQ ID NO: 580);  
 Ac-Val-Ile-Val-Ser-Ala-Arg-Met-Ala-(therapeutic agent) (SEQ ID NO: 581);  
 Ac-Val-Ile-Val-Ser-Ala-Arg-nLeu-Ala-(therapeutic agent) (SEQ ID NO: 582);
- 5 Ac-Val-Ser-Ala-Arg-nLeu-Ala-(therapeutic agent) (SEQ ID NO: 583);  
 Ac-Ile-Val-Ser-Ala-Arg-nLeu-Ala-(therapeutic agent) (SEQ ID NO: 584);  
 Ac-Gly-Ser-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 585);  
 Ac-Gly-Ser-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 586);  
 Ac-Gly-Ser-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 587);
- 10 Ac-Ser-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 588);  
 Ac-Ser-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 589);  
 Ac-Ser-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 590);  
 Ac-Arg-Gly-Ser-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 591);  
 Ac-Arg-Gly-Ser-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 592);
- 15 Ac-Arg-Gly-Ser-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 593);  
 Ac-Leu-Arg-Gly-Ser-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 594);  
 Ac-Leu-Arg-Gly-Ser-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 595); and
- 20 Ac-Leu-Arg-Gly-Ser-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 596).

In another embodiment, the conjugates provided herein are selected from:

- 25 Ac-R-Q-G-R-S-L-(therapeutic agent) (SEQ ID NO: 491);  
 Ac-R-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 492);  
 Ac-R-Q-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 493);  
 Ac-R-Q-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 494);  
 Ac-R-Q-G-R-S-F-(therapeutic agent) (SEQ ID NO: 495);



-124-

- Ac-R-Q-G-R-A-L-(therapeutic agent) (SEQ ID NO: 496);  
Ac-R-Q-G-R-A-L-(therapeutic agent) (SEQ ID NO: 497);  
Ac-R-Q-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 498);  
Ac-R-Q-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 499);  
5 Ac-R-Q-G-R-A-nV-(therapeutic agent) (SEQ ID NO: 500);  
Ac-R-Q-G-R-A-Cha-(therapeutic agent) (SEQ ID NO: 501);  
Ac-R-Q-G-R-A-F-(therapeutic agent) (SEQ ID NO: 502);  
Ac-R-N-G-R-S-L-(therapeutic agent) (SEQ ID NO: 503);  
Ac-R-N-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 504);  
10 Ac-R-Q-A-R-S-L-(therapeutic agent) (SEQ ID NO: 505);  
Ac-R-Q-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 506);  
Ac-R-Q-A-R-S-nV-(therapeutic agent) (SEQ ID NO: 507);  
Ac-R-Q-A-A-S-Cha-(therapeutic agent) (SEQ ID NO: 508);  
Ac-R-Q-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 509);  
15 Ac-R-Q-A-R-T-nL-(therapeutic agent) (SEQ ID NO: 510);  
Ac-R-Q-A-R-A-L-(therapeutic agent) (SEQ ID NO: 511);  
Ac-R-Q-A-R-A-nL-(therapeutic agent) (SEQ ID NO: 512);  
Ac-R-Q-A-R-A-nV-(therapeutic agent) (SEQ ID NO: 513);  
Ac-R-Q-A-R-A-Cha-(therapeutic agent) (SEQ ID NO: 514);  
20 Ac-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 515);  
Ac-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 516);  
Ac-R-Q-S-R-A-nL-(therapeutic agent) (SEQ ID NO: 517);  
Ac-R-Q-S-R-A-L-(therapeutic agent) (SEQ ID NO: 518);  
Ac-R-Q-S-R-A-nV-(therapeutic agent) (SEQ ID NO: 519);  
25 Ac-R-Q-S-R-A-Cha-(therapeutic agent) (SEQ ID NO: 520);  
Ac-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 521);  
Ac-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 522);  
Ac-R-Q-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 523);  
Ac-R-Q-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 524);

-125-

- Ac-R-Q-S-R-S-nV-(therapeutic agent) (SEQ ID NO: 525);  
Ac-R-Q-S-R-S-allylG-(therapeutic agent) (SEQ ID NO: 526);  
Ac-R-Q-S-R-S-Cha-(therapeutic agent) (SEQ ID NO: 527);  
Ac-R-Q-S-R-T-nL-(therapeutic agent) (SEQ ID NO: 528);  
5 Ac-R-Q-T-R-S-S-L-(therapeutic agent) (SEQ ID NO: 529);  
Ac-R-Q-T-R-S-L-(therapeutic agent) (SEQ ID NO: 530);  
Ac-R-N-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 531);  
Ac-R-Q-F-R-S-L-(therapeutic agent) (SEQ ID NO: 532);  
Ac-R-Q-F-R-S-nL-(therapeutic agent) (SEQ ID NO: 534);  
10 Ac-R-Q-F-R-S-nV-(therapeutic agent) (SEQ ID NO: 535);  
Ac-R-Q-F-R-S-nL-(therapeutic agent) (SEQ ID NO: 536);  
Ac-R-Q-F-R-S-Cha-(therapeutic agent) (SEQ ID NO: 537);  
Ac-R-Q-F-R-A-L-(therapeutic agent) (SEQ ID NO: 538);  
Ac-R-Q-F-R-A-nL-(therapeutic agent) (SEQ ID NO: 539);  
15 Ac-R-Q-F-R-A-nV-(therapeutic agent) (SEQ ID NO: 540);  
Ac-R-Q-F-R-A-Cha-(therapeutic agent) (SEQ ID NO: 541);  
Ac-Q-S-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 542);  
MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 483);  
MeOCO-Quat3-G-R-S-L-(therapeutic agent) (SEQ ID NO: 484);  
20 MeOCO-Quat-G-R-S-L-(therapeutic agent) (SEQ ID NO: 485);  
MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 486);  
MeOCO-Quat5-G-R-S-L-(therapeutic agent) (SEQ ID NO: 487);  
MeOCO-Quat2-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 488);  
MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 489);  
25 MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 490);  
Ac-Q-G-R-S-L-(therapeutic agent) (SEQ ID NO: 445);  
Ac-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 446);  
Ac-Q-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 447);  
Ac-N-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 448);

-126-

- Ac-Q-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 449);  
Ac-Q-G-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 450);  
Ac-Q-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 451);  
Ac-Q-G-R-S-S-allylG-(therapeutic agent) (SEQ ID NO: 452);  
5 Ac-Q-G-R-S-S-allylG-(therapeutic agent) (SEQ ID NO: 453);  
Ac-Q-A-R-S-L-(therapeutic agent) (SEQ ID NO: 454);  
Ac-Q-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 455);  
Ac-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 456);  
Ac-Q-S-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 457);  
10 Ac-Q-S-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 458);  
Ac-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 459);  
Ac-Q-T-R-S-S-L-(therapeutic agent) (SEQ ID NO: 460);  
Ac-Q-Aib-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 461);  
Ac-Q-Aib -R-S-S-L-(therapeutic agent) (SEQ ID NO: 462);  
15 Ac-Q-Abu-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 463);  
Ac-Q-Abu-R-S-S-L-(therapeutic agent) (SEQ ID NO: 464);  
Ac-Q-Cha-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 465);  
Ac-Q-F-R-S-L-(therapeutic agent) (SEQ ID NO: 466);  
Ac-Q-F-R-S-S-L-(therapeutic agent) (SEQ ID NO: 467);  
20 Ac-Q-Y-R-S-S-L-(therapeutic agent) (SEQ ID NO: 468);  
Ac-R-G-R-S-L-(therapeutic agent) (SEQ ID NO: 469);  
Ac-R-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 470);  
Ac-R-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 471);  
Ac-R-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 472);  
25 Ac-R-A-R-S-L-(therapeutic agent) (SEQ ID NO: 473);  
Ac-R-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 474);  
Ac-R-S-R-S-L-(therapeutic agent) (SEQ ID NO: 475);  
Ac-R-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 476);  
Ac-R-S-R-S-Cha-(therapeutic agent) (SEQ ID NO: 477);

-127-

- Ac-R-S-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 478);
- Ac-R-F-R-S-L-(therapeutic agent) (SEQ ID NO: 479);
- Ac-R-F-R-S-Cha-(therapeutic agent) (SEQ ID NO: 480);
- Ac-Y-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 481);
- 5 Ac-M(O2)-S-R-S-L-(therapeutic agent) (SEQ ID NO: 482);
- Ac-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 105);
- Ac-R-R-Q-S-R-I-(therapeutic agent) (SEQ ID NO: 610);
- Ac-R-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 543);
- Ac-R-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 544);
- 10 Ac-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 545);
- Ac-R-G-S-G-R--S-nL-(therapeutic agent) (SEQ ID NO: 546);
- Ac-R-G-S-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 547);
- Ac-R-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 548);
- Ac-I-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 549);
- 15 Ac-R-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 108);
- Ac-R-R-Q-S-R-I-(therapeutic agent) (SEQ ID NO: 111);
- Ac-L-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 106);
- Ac-L-R-R-Q-S-R-G-G-(therapeutic agent) (SEQ ID NO: 109);
- Ac-L-R-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 110);
- 20 Ac-L-R-R-Q-S-R-A-I-(therapeutic agent) (SEQ ID NO: 112);
- Ac-L-R-R-Q-S-R-A-I-(therapeutic agent) (SEQ ID NO: 611);
- Ac-L-R-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 550); and
- Ac-L-R-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 551);

In another embodiment, the conjugates provided herein are  
 25 selected from:

- Ac-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 362);
- Ac-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 363);
- Ac-S-G-R-S-S-S-L-(therapeutic agent) (SEQ ID NO: 364);
- Ac-S-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 365);

-128-

- Ac-S-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 366); isomer 1  
Ac-S-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 367); isomer 2  
Ac-S-G-R-S-G(hex)-(therapeutic agent) (SEQ ID NO: 368);  
Ac-S-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 369);  
5 Ac-S-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 370);  
Ac-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 371);  
Ac-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 372);  
Ac-S-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 373);  
Ac-T-G-R-S-Abu-(therapeutic agent) (SEQ ID NO: 374);  
10 Ac-T-G-R-S-L-(therapeutic agent) (SEQ ID NO: 375);  
Ac-T-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 376);  
Ac-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 377);  
Ac-T-G-R-S-G(hex)-(therapeutic agent) (SEQ ID NO: 378);  
Ac-T-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 379);  
15 Ac-T-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 380);  
Ac-T-G-R-T-Abu-(therapeutic agent) (SEQ ID NO: 381);  
Ac-T-G-R-hS-nL-(therapeutic agent) (SEQ ID NO: 382);  
Ac-T-G-R-Abu-nL-(therapeutic agent) (SEQ ID NO: 383);  
Ac-T-G-R-Abu-nV-(therapeutic agent) (SEQ ID NO: 384);  
20 Ac-T-G-F(Gn)-S-nL-(therapeutic agent) (SEQ ID NO: 385);  
Ac-T-G-F(Gn)-S-Cha-(therapeutic agent) (SEQ ID NO: 386);  
Ac-T-G-F(Gn)-Abu-nV-(therapeutic agent) (SEQ ID NO: 387);  
Ac-T-G-K(alloc)-S-nL-(therapeutic agent) (SEQ ID NO: 388);  
Ac-T-G-K-S-nL-(therapeutic agent) (SEQ ID NO: 389);  
25 Ac-T-G-hR-S-nL-(therapeutic agent) (SEQ ID NO: 390);  
Ac-(hS)G-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 391);  
MeOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 392);  
PhSO<sub>2</sub>-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 393);  
MeOEtCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 394);

-129-

- MeO(EtO)<sub>2</sub>Ac-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 395);  
4-oxo-Pentanoyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 396);  
3,4-MethyldioxyPhAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 397);  
2-PyridylAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 398);
- 5 PhOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 399);  
L-3-PhLactyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 400);  
MeOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 401);  
PhAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 402);  
MeOEtOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 403);
- 10 MeOEtOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 404);  
HOOCButa-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 405);  
Z-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 406);  
EtOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 407);  
 $\beta$ A-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 408);
- 15 Pent-4-ynoyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 409);  
NapAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 410);  
iBoc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 411);  
HOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 412);  
MeSucc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 413);
- 20 N,N-diMeGly-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 414);  
Succ-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 415);  
HCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 416);  
Ac-T-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 417);  
Ac-T-A-F(Gn)-S-nL-(therapeutic agent) (SEQ ID NO: 418);
- 25 Ac-T-A-R-Abu-nV-(therapeutic agent) (SEQ ID NO: 419);  
Ac-T-A-R-S-Abu-(therapeutic agent) (SEQ ID NO: 420);  
Ac-T-A-R-T-Abu-(therapeutic agent) (SEQ ID NO: 421);  
Ac-T-S(O-Me)-R-S-nL-(therapeutic agent) (SEQ ID NO: 422);  
Ac-T-hS-R-S-nL-(therapeutic agent) (SEQ ID NO: 423);

-130-

- Ac-T-(1-Me)H-R-S-nL-(therapeutic agent) (SEQ ID NO: 424);  
Ac-T-(3-Me)H-R-S-nL-(therapeutic agent) (SEQ ID NO: 425);  
Ac-T-H-R-S-nL-(therapeutic agent) (SEQ ID NO: 426);  
Ac-T-Sar-R-S-nL-(therapeutic agent) (SEQ ID NO: 427);  
5 Ac-T-nV-R-S-nL-(therapeutic agent) (SEQ ID NO: 428);  
Ac-T-nL-R-S-nL-(therapeutic agent) (SEQ ID NO: 429);  
Ac-T-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 430);  
Ac-T-Abu-R-S-nL-(therapeutic agent) (SEQ ID NO: 431);  
Ac-4,4diMeThr-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 432);  
10 Ac-hS-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 433);  
Ac-hS-G-R-hS-Cha-(therapeutic agent) (SEQ ID NO: 434);  
Ac-hS-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 435);  
Ac-hS-G-R-T-Cha-(therapeutic agent) (SEQ ID NO: 436);  
Ac-hS-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 437);  
15 Ac-N-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 438);  
Ac-Y-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 439);  
Ac-Y-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 440);  
Ac-Q-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 441);  
Ac-Q-G-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 442);  
20 Ac-L-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 573);  
Ac-L-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 342);  
Ac-L-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 343);  
Ac-L-R-G-S-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 344);  
Ac-L-R-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 345);  
25 Ac-L-R-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 574);  
Ac-L-R-G-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 346);  
Ac-L-R-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 347);  
Ac-L-R-G-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 348);  
Ac-L-R-G-S-A-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 349);

-131-

- Ac-L-R-G-S-A-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 350);  
Ac-V-I-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 351);  
Ac-V-I-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 352);  
Ac-V-I-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 353);  
5 Ac-V-I-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 354);  
Ac-V-I-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 355);  
Ac-V-I-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 356);  
Ac-V-I-V-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 357);  
Ac-V-I-V-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 358);  
10 Ac-V-I-V-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 359);  
Ac-R-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 360);  
Ac-R-R-nV-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 361);  
Ac-R-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 309);  
Ac-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 310);  
15 Ac-R-G-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 311);  
Ac-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 312);  
Ac-R-G-S-G-R--S-nL-(therapeutic agent) (SEQ ID NO: 313);  
Ac-R-G-S-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 314);  
Ac-R-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 315);  
20 Ac-R-G-S-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 316);  
Ac-R-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 317);  
Ac-R-G-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 318);  
Ac-R-G-S-A-R-S-S-(therapeutic agent) (SEQ ID NO: 319);  
Ac-R-G-S-A-R-S-nV-(therapeutic agent) (SEQ ID NO: 320);  
25 Ac-R-G-S-A-R-S-S-nV -(therapeutic agent) (SEQ ID NO: 321);  
Ac-R-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 322);  
Ac-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 323);  
Ac-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 324);  
Ac-R-C(Me)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 325);



-132-

- Ac-R-L-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 326);  
Ac-R-V-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 327);  
Ac-R-V-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 328);  
Ac-R-nL-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 329);  
5 Ac-R-G(tBu)-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 330);  
Ac-R-L-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 331);  
Ac-R-V-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 332);  
Ac-R-nL-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 333);  
Ac-I-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 334);  
10 Ac-I-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 335);  
Ac-I-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 336);  
Ac-I-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 337);  
Ac-I-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 338);  
Ac-I-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 339);  
15 Ac-I-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 340);  
Ac-I-V-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 341);  
Ac-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 585);  
Ac-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 277);  
Ac-G-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 278);  
20 Ac-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 279);  
Ac-G-S-G-R-L-(therapeutic agent) (SEQ ID NO: 280);  
Ac-G-S-G-(4-guan)Phg-S-L-(therapeutic agent) (SEQ ID NO: 281);  
Ac-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 282);  
Ac-G-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 283);  
25 Ac-G-S-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 284);  
Ac-G-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 285);  
Succ-bA-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 286);  
Ac-G-T-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 287);  
Ac-G-hS-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 288);

-133-

- Ac-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 289);
- Ac-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 290);
- Ac-G-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 291);
- Ac-G-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 292);
- 5 Ac-G-S-A-R-A-S-L-(therapeutic agent) (SEQ ID NO: 293);
- Ac-V-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 294);
- Ac-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 295);
- Ac-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 296);
- Ac-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 297);
- 10 Ac-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 298);
- Ac-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 299);
- Ac-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 300);
- Ac-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 301);
- Ac-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 302);
- 15 Ac-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 303);
- Ac-C(Me)-P-G-R-A-L-(therapeutic agent) (SEQ ID NO: 304);
- Ac-C(Me)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 305);
- Ac-C(Me)-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 306);
- Ac-C(Me)-P-A-R-A-S-L-(therapeutic agent) (SEQ ID NO: 307);and
- 20 Ac-G(tBu)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 308);

In another embodiment, the conjugates provided herein are selected from:

- Ac-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 552);
- Ac-Q-S-R-S-A-(therapeutic agent) (SEQ ID NO: 553);
- 25 Ac-Q-S-R-S-G-(therapeutic agent) (SEQ ID NO: 554);
- Ac-R-S-R-A-A-(therapeutic agent) (SEQ ID NO: 555);
- Ac-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 556);
- Ac-R-Q-S-R-S-A-(therapeutic agent) (SEQ ID NO: 557); and
- Ac-R-Q-S-R-S-A-A-(therapeutic agent) (SEQ ID NO: 558);

-134-

In another embodiment, the conjugates provided herein are selected from:

- Ac-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 559);
- Ac-S-G-R-A-A-(therapeutic agent) (SEQ ID NO: 560);
- 5 Ac-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 561);
- Ac-S-G-R-S-S-A-(therapeutic agent) (SEQ ID NO: 562);
- Ac-S-G-R-A-S-A-(therapeutic agent) (SEQ ID NO: 563);
- Ac-S-G-R-S-G-(therapeutic agent) (SEQ ID NO: 564);
- Ac-S-G-R-S-S-G-(therapeutic agent) (SEQ ID NO: 565);
- 10 Ac-S-G-R-S-G-A-(therapeutic agent) (SEQ ID NO: 566);
- Ac-S-G-R-S-G-G-(therapeutic agent) (SEQ ID NO: 567); and
- Ac-G-T-G-R-S-G-G-(therapeutic agent) (SEQ ID NO: 568);

In another embodiment, the conjugates provided herein are selected from:

- 15 Ac-L-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 597);
- MeSO<sub>2</sub>-dA(Chx)-Abu-R-S-L-(therapeutic agent) (SEQ ID NO: 598);
- Ac-R-A-R-S-L-(therapeutic agent) (SEQ ID NO: 599);
- Ac-dA(Chx)-Abu-R-S-L-(therapeutic agent) (SEQ ID NO: 600);
- Ac-dA(Chx)-Abu-R-S-S-L-(therapeutic agent) (SEQ ID NO: 601);
- 20 Ac-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 602);
- MeOCO-dhF-P(OH)-R-S-S-L-(therapeutic agent) (SEQ ID NO: 603);
- MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 604);
- Ac-dCha-P(OH)-R-S-S-L-(therapeutic agent) (SEQ ID NO: 605);
- Ac-dCha-Abu-R-S-S-A-(therapeutic agent) (SEQ ID NO: 606);
- 25 MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 607);
- MeOCO-Quat3-G-R-S-L-(therapeutic agent) (SEQ ID NO: 608); and
- MeOCO-Quat-G-R-S-L-(therapeutic agent) (SEQ ID NO: 609).

It also understood that conjugates containing the above peptidic substrate portions can be prepared with other capping groups in place of

-135-

Ac (see, *e.g.*, the description herein of the capping groups X<sup>n</sup>).

Therapeutic agents for use in the above conjugates include, for example, cytotoxic agents, such as, but not limited to, a toxin such as abrin, ricin A, pseudomonas exotoxin shiga toxin, diphtheria toxin and other  
5 such toxins and toxic portions thereof; proteins such as tumor necrosis factor, interferons, such as  $\alpha$ -interferon and gamma-interferon, pro-coagulants such as tissue factor and tissue factor variants, pro-apoptotic agents such FAS-ligand, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; biological response modifiers such  
10 as, for example, lymphokines, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), fibroblast growth factors (FGFs) and other growth factors, the methotrexate group of drugs, the anthracycline family of drugs, the vinca alkaloid drugs, the  
15 mitomycins, the bleomycins, the cytotoxic nucleosides including cytosine arabinosides and difluoronucleosides, the pteridine family of drugs, diynenes, the taxanes and the podophyllotoxins. All such conjugates are within the scope of the instant disclosure and can be prepared and used as disclosed herein.

20 Thus, the conjugates provided herein include, but are not limited to, those disclosed herein where the therapeutic agent is, *e.g.*, doxorubicin, carminomycin, daunorubicin, detorubicin, idarubicin, epirubicin, esorubicin, THP, AD-32, aminopterin, methotrexate, methopterin, dichloromethotrexate, mitomycin C, porfiromycin,  
25 5-fluorouracil, 6-mercaptopurine, cytosine arabinoside, podophyllotoxin, or podophyllotoxin derivatives such as etoposide or etoposide phosphate, melphalan, vinblastine, vincristine, leurosidine, vindesine, leurosine, taxol, estramustine, cisplatin, combretastatin and analogs, and

-136-

cyclophosphamide. In one embodiment, the therapeutic agent is doxorubicin. In another embodiment, the therapeutic agent is taxol.

Any conjugates corresponding to the above conjugates or any conjugates disclosed herein where the P1' and/or P2' residues are Ile in  
5 place of Ala are within the scope of the instant disclosure and can be prepared and used as disclosed herein.

Any peptidic substrates formed by permutation and selection of amino acids from those set forth in the above definitions of P groups are contemplated.

#### 10 D. Preparation of the Conjugates

The peptidic substrates of the conjugates provided herein are synthesized from their constituent amino acids by conventional peptide synthesis techniques, such as by solid-phase technology. The peptidic substrates are then purified by reverse-phase high performance liquid  
15 chromatography (HPLC).

The peptide acids can be prepared from their constituent Fmoc-aminoacids. Standard methods of peptide synthesis are disclosed, for example, in the following works: Synthesis Notes Section, NovaBiochem Catalog 2002/3, Schroeder *et al.*, "The Peptides", Vol. 1, Academic  
20 Press 1965; Bodansky *et al.*, "Peptide Synthesis", Interscience Publishers, 1966; McOmie (ed.) "Protective Groups in Organic Chemistry", Plenum Press, 1973, Barany *et al.*, "The Peptides: Analysis, Synthesis, Biology" 2, Chapter 1, Academic Press, 1990, and Stewart *et al.*, "Solid Phase Peptide Synthesis", Second Edition, Pierce Chemical  
25 Company, 1994. The disclosures of these references are hereby incorporated by reference.

The pharmaceutically acceptable salts of the conjugates provided herein include the conventional non-toxic salts of the conjugates as formed, *e.g.*, from non-toxic inorganic or organic acids. For example,

-137-

such conventional non-toxic salts include those derived from inorganic acids such as hydrochloric, hydrobromic, sulfuric, sulfamic, phosphoric, nitric and the like: and the salts prepared from organic acids such as acetic, propionic, succinic, glycolic, stearic, lactic, malic, tartaric, citric, 5 ascorbic, pantoic, maleic, hydroxymaleic, phenyl-acetic, glutamic, benzoic, salicylic, sulfanilic, 2-acetoxy-benzoic, fumaric, toluenesulfonic, methanesulfonic, ethane disulfonic, oxalic, isethionic, trifluoroacetic and the like.

The conjugates provided herein that contain the peptidic moieties 10 containing the cell surface protease cleavage site and a therapeutic agent can similarly be synthesized by techniques known to those of skill in the art. For example, a free amine moiety on the therapeutic agent can be covalently attached to the peptidic substrate at the carboxyl terminus such that an amide bond is formed. Similarly, an amide bond can be 15 formed by covalently coupling an amine moiety of the peptidic substrate and a carboxyl moiety of the therapeutic agent. For these purposes a reagent such as 2-(1H-benzotriazol-1-yl)-1,3,3-tetramethyl-uronium hexafluorophosphate (known as HBTU) and 1-hydroxybenzotriazole hydrate (known as HOBT), dicyclohexyl-carbodiimide (DCC), 20 N-ethyl-N-(3-dimethylaminopropyl)-carbodiimide (EDC), diphenyl-phosphorylazide (DPPA), benzotriazol-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate (BOP) and the like, used in combination or singularly, can be utilized.

Furthermore, the instant conjugates can be formed by a non- 25 peptidyl bond between the cell surface protease cleavage site and a therapeutic agent. For example, the therapeutic agent can be covalently attached to the carboxyl terminus of the peptidic substrate via a hydroxyl moiety on the therapeutic agent, thereby forming an ester linkage. For this purpose a reagent such as a combination of HBTU and HOBT, a

-138-

combination of BOP and imidazole, a combination of DCC and DMAP, and the like can be utilized. The carboxylic acid also can be activated by forming the nitro-phenyl ester or the like and reacted in the presence of DBU (1,8-diazabicyclo[5,4,0]undec-7-ene).

- 5           The instant conjugates also can be formed by attachment of the peptidic substrate to the therapeutic agent via a linker unit. Such linker units include, for example, a biscarbonyl alkyl diradical whereby an amine moiety on the therapeutic agent is connected with the linker unit to form an amide bond and the amino terminus of the peptidic substrate is
- 10   connected with the other end of the linker unit also forming an amide bond. Conversely, a diaminoalkyl diradical linker unit, whereby a carbonyl moiety on the cytotoxic agent is covalently attached to one of the amines of the linker unit while the other amine of the linker unit is covalently attached to the C-terminus of the peptidic substrate, also can be useful.
- 15   Other such linker units which are stable to the physiological environment when not in the presence of a cell surface protease, or a soluble, shed or released form thereof, but are cleavable upon the cleavage of the cell surface protease proteolytic cleavage site, or a soluble, shed or released form thereof, are also envisioned. Furthermore, linker units can be utilized
- 20   that, upon cleavage of the cell surface protease proteolytic cleavage site, remain attached to the therapeutic agent but do not significantly decrease the therapeutic activity of such a post-cleavage therapeutic agent derivative when compared with an unmodified therapeutic agent.

- One skilled in the art understands that in the synthesis of the
- 25   conjugates provided herein, one can need to protect various reactive functionalities on the starting compounds and intermediates while a desired reaction is carried out on other portions of the molecule. After the desired reactions are complete, or at any desired time, normally such protecting groups will be removed by, for example, hydrolytic or

-139-

hydrogenolytic means. Such protection and deprotection steps are conventional in organic chemistry. One skilled in the art is referred to Protective Groups in Organic Chemistry, McOmie, ed., Plenum Press, NY, NY (1973); and, Protective Groups in Organic Synthesis, Greene, ed.,  
5 John Wiley & Sons, NY, NY (1991) for the teaching of protective groups which can be useful in the preparation of the conjugates provided herein.

By way of example only, useful amino-protecting groups can include, for example, C<sub>1</sub>-C<sub>10</sub> alkanoyl groups such as formyl, acetyl, dichloroacetyl, propionyl, hexanoyl, 3,3-diethylhexanoyl,  $\gamma$ -chlorobutryl,  
10 and the like; C<sub>1</sub>-C<sub>10</sub> alkoxy carbonyl and C<sub>5</sub>-C<sub>15</sub> aryloxy carbonyl groups such as tert-butoxy carbonyl, benzyloxy carbonyl, allyloxy carbonyl, 4-nitrobenzyloxy carbonyl, fluorenylmethyloxy carbonyl and cinnamoyloxy carbonyl; halo(C<sub>1</sub>-C<sub>10</sub>)-alkoxy carbonyl such as 2,2,2-trichloroethoxy carbonyl; and C<sub>1</sub>-C<sub>15</sub> arylalkyl and alkenyl group such as benzyl,  
15 phenethyl, allyl, trityl, and the like. Other commonly used amino-protecting groups are those in the form of enamines prepared with  $\beta$ -keto-esters such as methyl or ethyl acetoacetate.

Useful carboxy-protecting groups can include, for example, C<sub>1</sub>-C<sub>10</sub> alkyl groups such as methyl, tert-butyl, decyl; halo C<sub>1</sub>-C<sub>10</sub> alkyl such as  
20 2,2,2-trichloroethyl, and 2-iodoethyl; C<sub>5</sub>-C<sub>15</sub> arylalkyl such as benzyl, 4-methoxybenzyl, 4-nitrobenzyl, triphenylmethyl, diphenylmethyl; C<sub>1</sub>-C<sub>10</sub> alkanoyloxymethyl such as acetoxy-methyl, propionoxymethyl and the like; and groups such as phenacyl, 4-halophenacyl, allyl, dimethylallyl, tri-(C<sub>1</sub>-C<sub>3</sub> alkyl)silyl, such as trimethylsilyl,  $\beta$ -p-toluenesulfonylethyl,  
25  $\beta$ -p-nitrophenyl-thioethyl, 2,4,6-trimethylbenzyl,  $\beta$ -methylthioethyl, phthalimidomethyl, 2,4-dinitro-phenylsulphenyl, 2-nitrobenzhydryl and related groups.

Similarly, useful hydroxy protecting groups can include, for example, the formyl group, the chloroacetyl group, the benzyl group, the



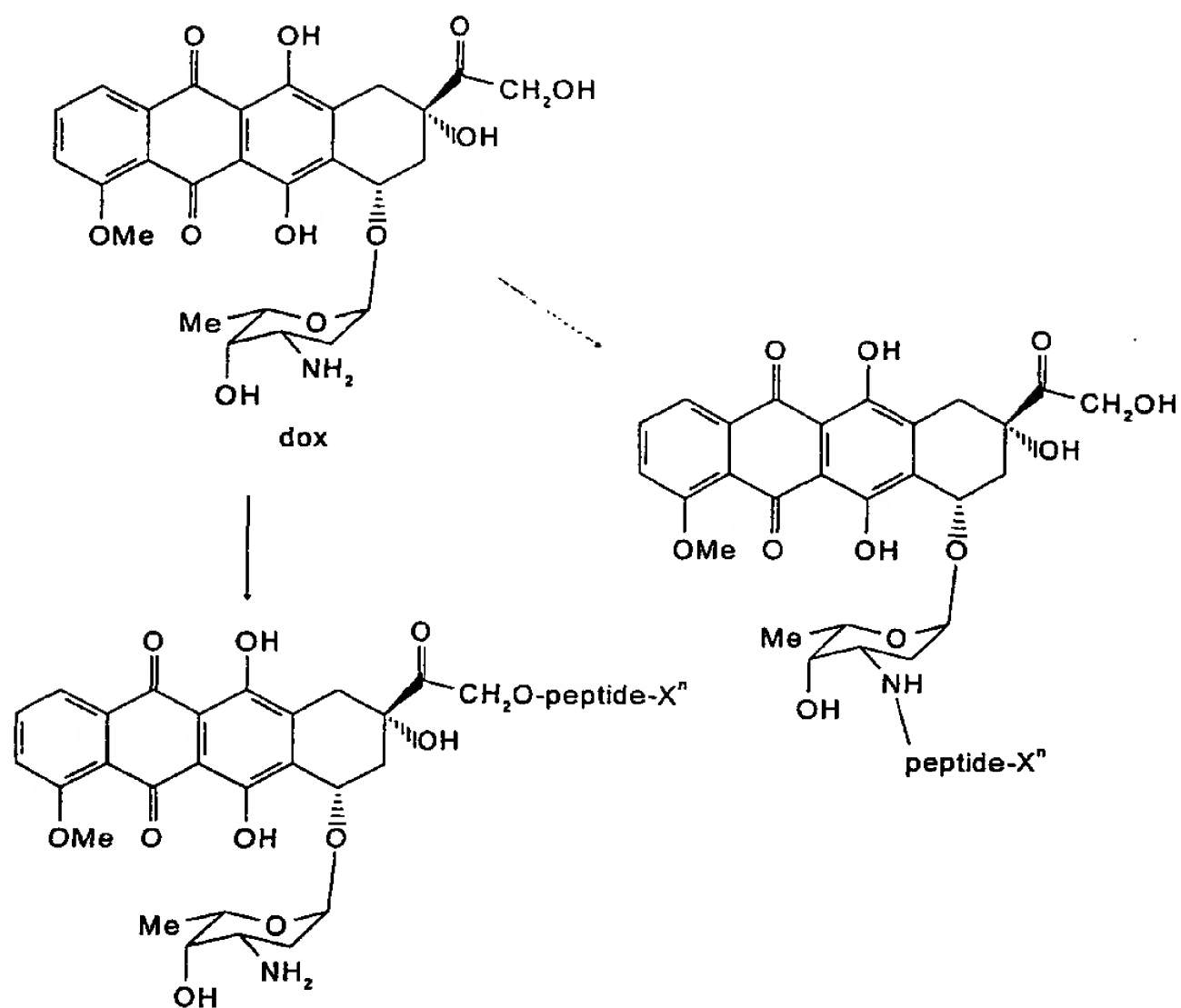
-140-

benzhydryl group, the trityl group, the 4-nitrobenzyl group, the trimethylsilyl group, the phenacyl group, the tert-butyl group, the methoxymethyl group, the tetrahydropyranyl group, the tert-butyl-dimethylsilyl group and the like.

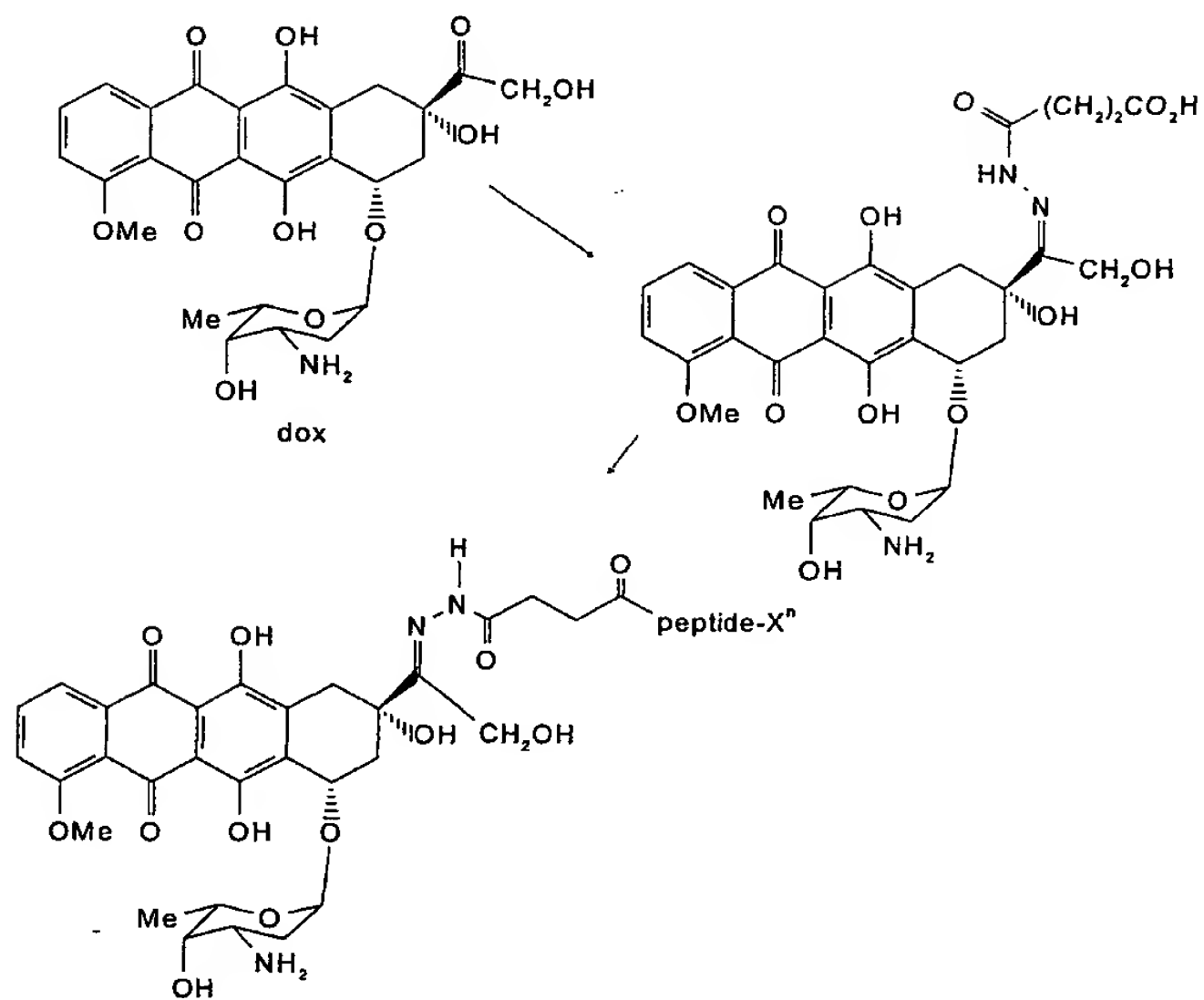
- 5           With respect to the embodiment of a peptidic substrate combined with the anthracycline antibiotic doxorubicin, the following Reaction Schemes illustrate the synthesis of the conjugates provided herein.

-141-

## REACTION SCHEME I

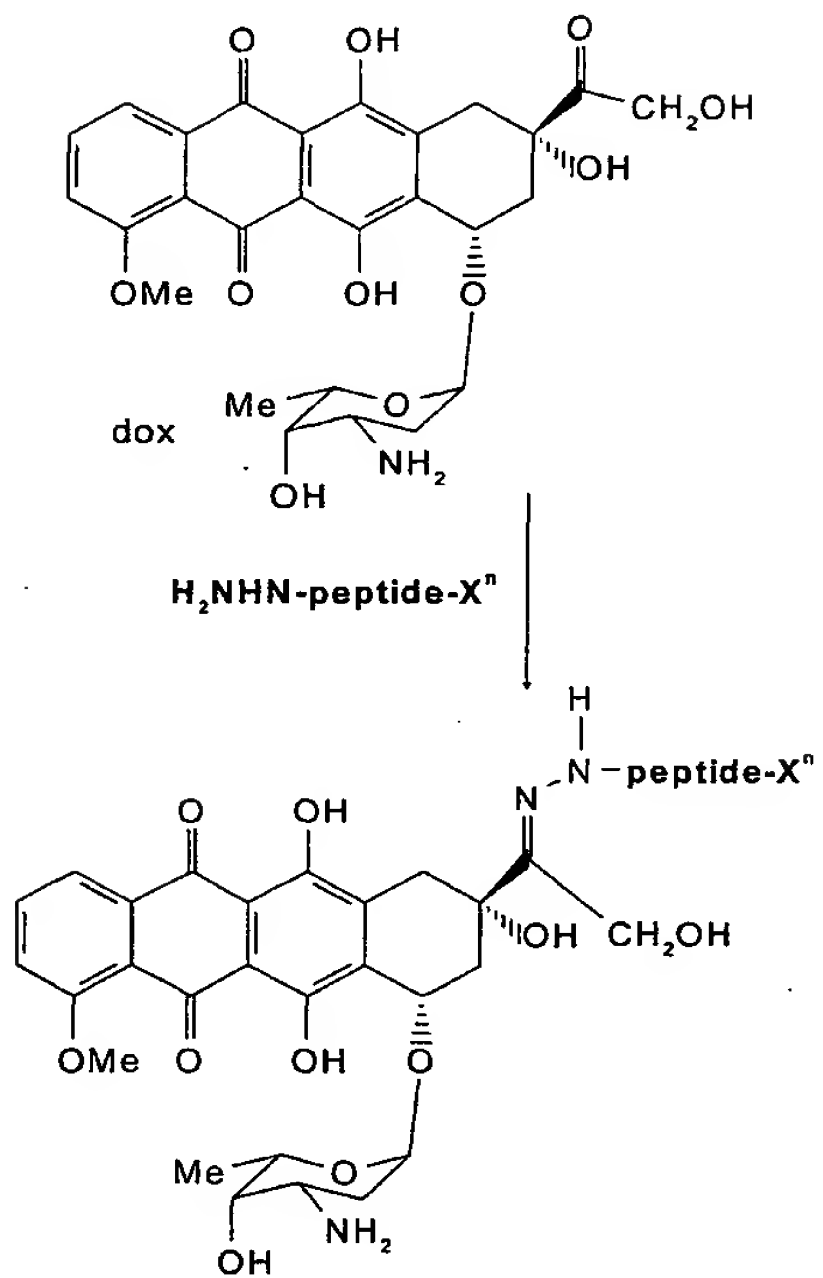






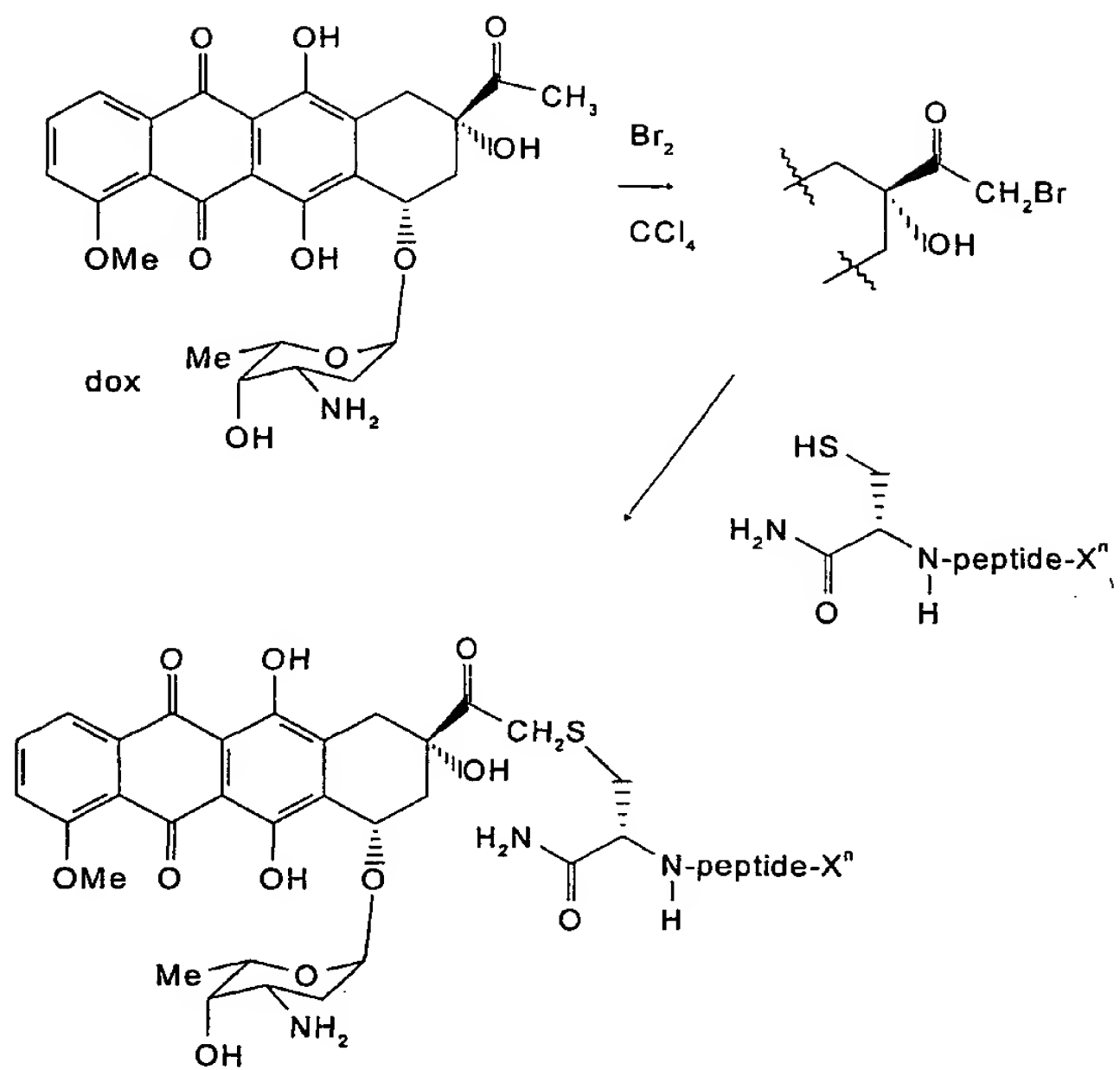
-144-

## REACTION SCHEME IV



-145-

## REACTION SCHEME V



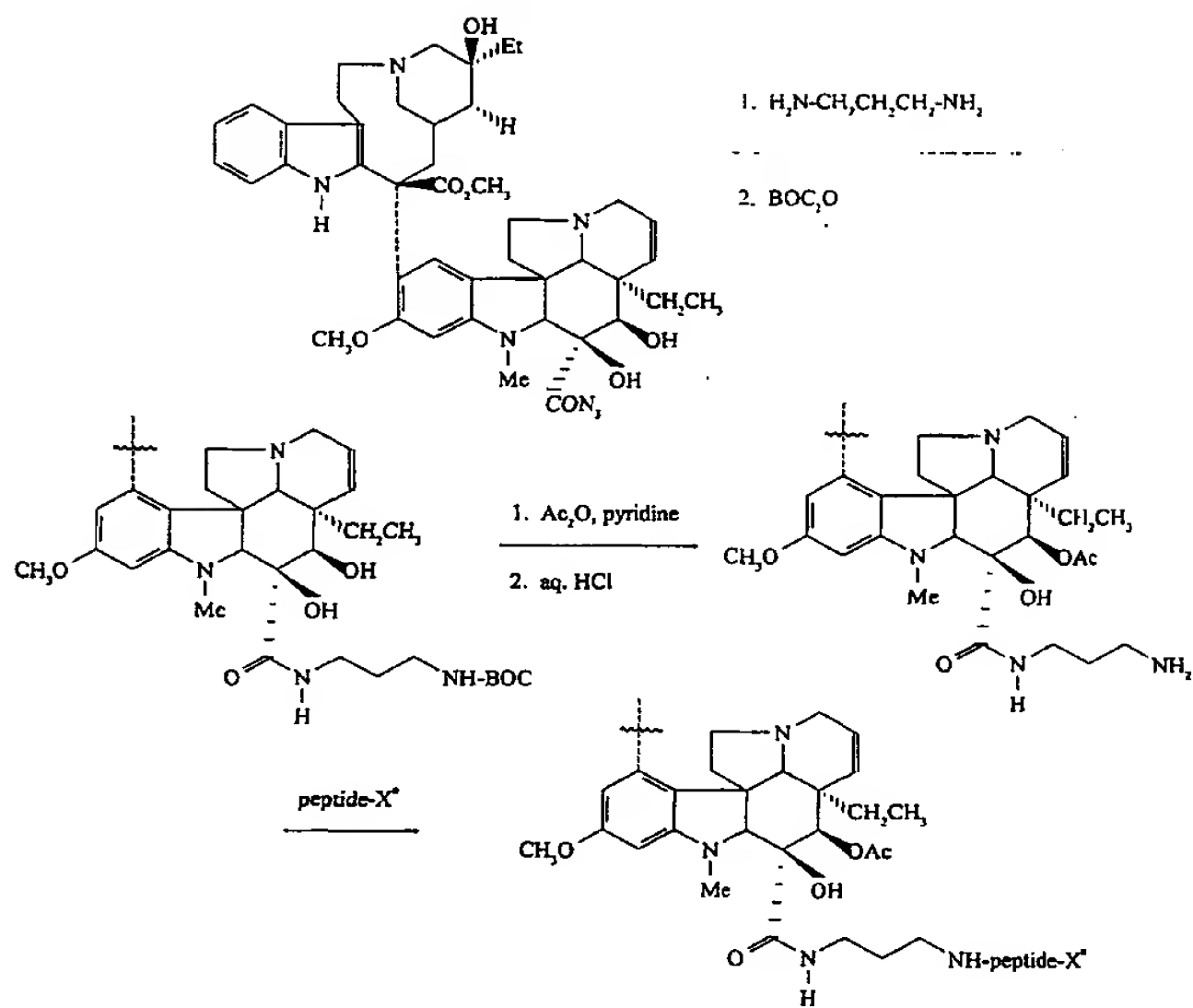
-146-

Reaction Scheme VI illustrates preparation of the conjugates provided herein of a peptidic substrate and the vinca alkaloid cytotoxic agent vinblastine wherein the attachment of vinblastine is at the C-terminus of the peptidic substrate. The use of the 1,3-diaminopropane linker is illustrative only; other linker units between the carbonyl of vinblastine and the C-terminus of the peptidic substrate are also envisioned (*e.g.*,  $(\text{CH}_2)_n\text{-T-(CH}_2)_n$ ). The acyl azide starting material is prepared from vinglasine by reaction with hydrazine (60-65 °C, MeOH), followed by reaction with HCl/DMF/isoamyl nitrite. Furthermore, Reaction Scheme VI illustrates a synthesis of conjugates wherein the C4-hydroxy moiety is reacylated following the addition of the linker unit. It is known that the desacetyl vinblastine conjugate also is efficacious and can be prepared by eliminating the steps shown in Reaction Scheme VI of protecting the primary amine of the linker and reacting the intermediate with acetic anhydride, followed by deprotection of the amine (see, *e.g.*, International Patent Application Publication No. WO 98/10651).

Conjugation of the peptidic substrate at other positions and functional groups of vinblastine can be readily accomplished by one of ordinary skill in the art and also is expected to provide conjugates that are substrates for cell surface proteases, or a soluble, shed or released form thereof.

-147-

## REACTION SCHEME VI





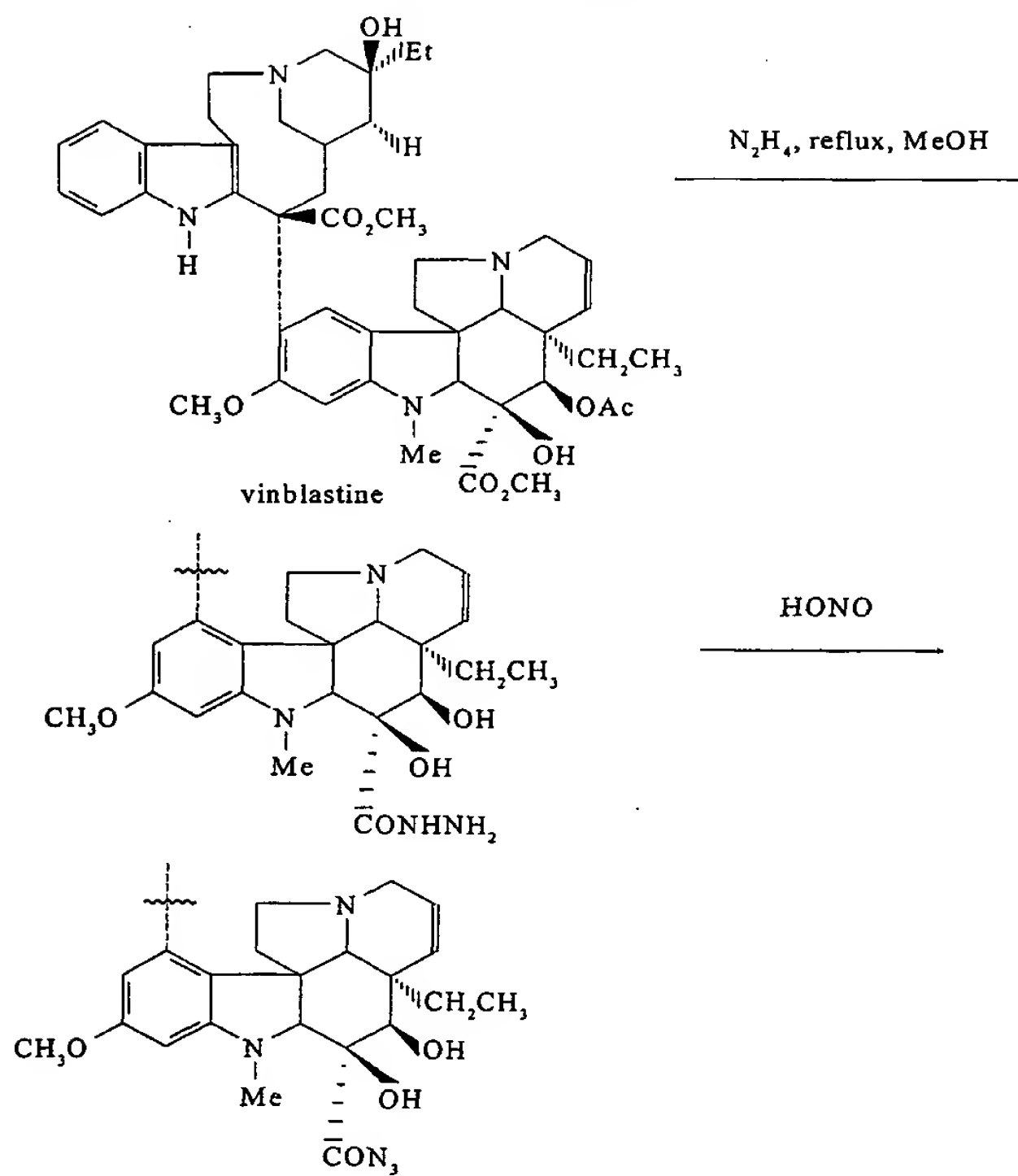
-148-

Reaction Scheme VII illustrates preparation of certain of the conjugates utilized in the compositions and methods provided herein wherein the peptidic substrates are combined with the vinca alkaloid cytotoxic agent vinblastine. Attachment of the N-terminus of the peptidic substrate to vinblastine is illustrated (S.P. Kandukuri *et al.* (1985) *J. Med. Chem.* 28:1079-1088).

It also is understood that conjugates can be prepared wherein the N-terminus of the peptidic substrate utilized in the compositions and methods provided herein is combined with one therapeutic agent, such as a cytotoxic agent, such as vinblastine, while the C-terminus is simultaneously attached to another cytotoxic agent, which is the same or different cytotoxic agent, such as doxorubicin. Reaction Scheme VIII illustrates the synthesis of such a polycytotoxic agent conjugate. Such a polycytotoxic conjugate can offer advantages over a conjugate containing only one cytotoxic agent.

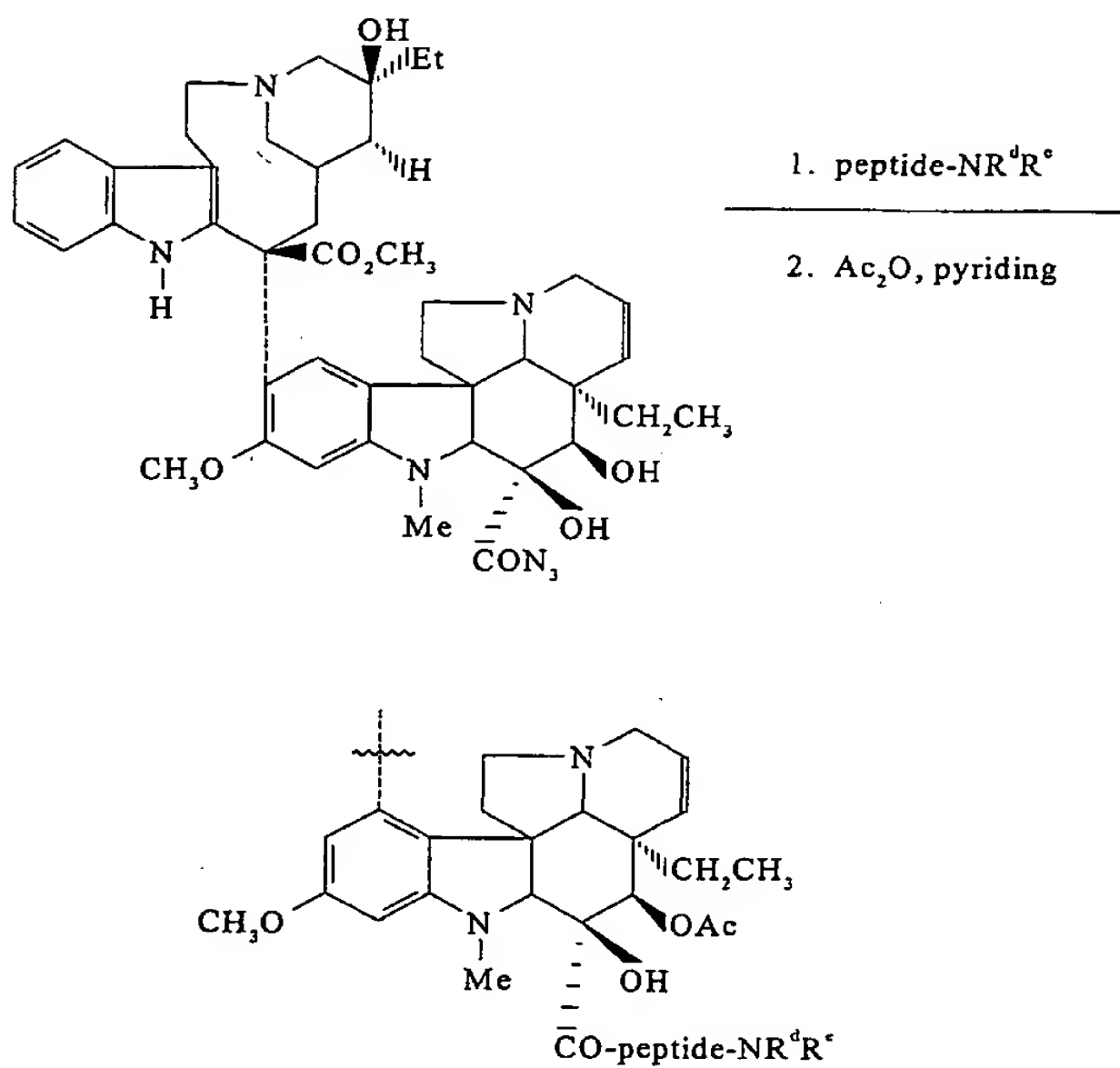
-149-

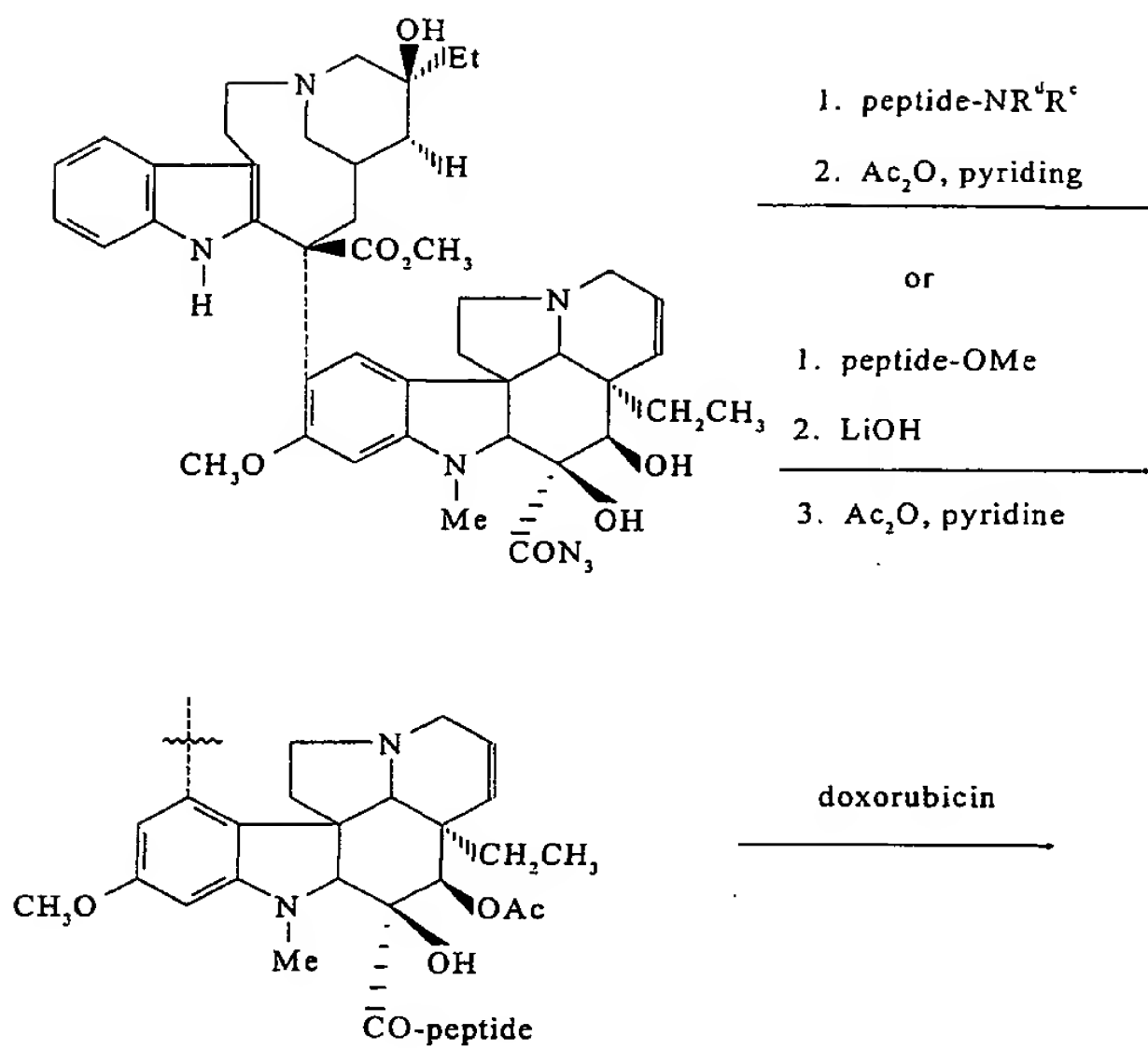
## REACTION SCHEME VII



-150-

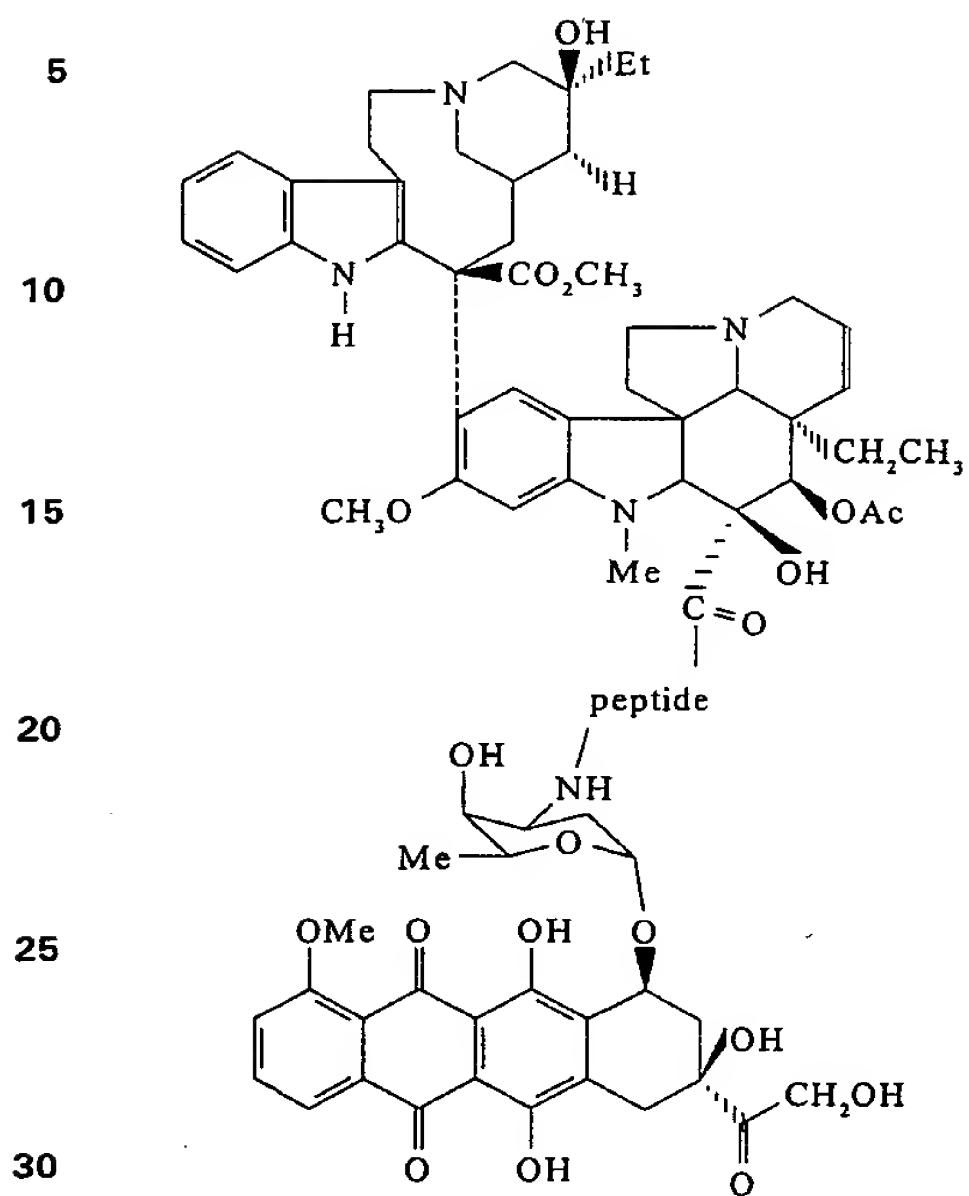
## REACTION SCHEME VII (Continued)





-152-

## REACTION SCHEME VIII (Continued)



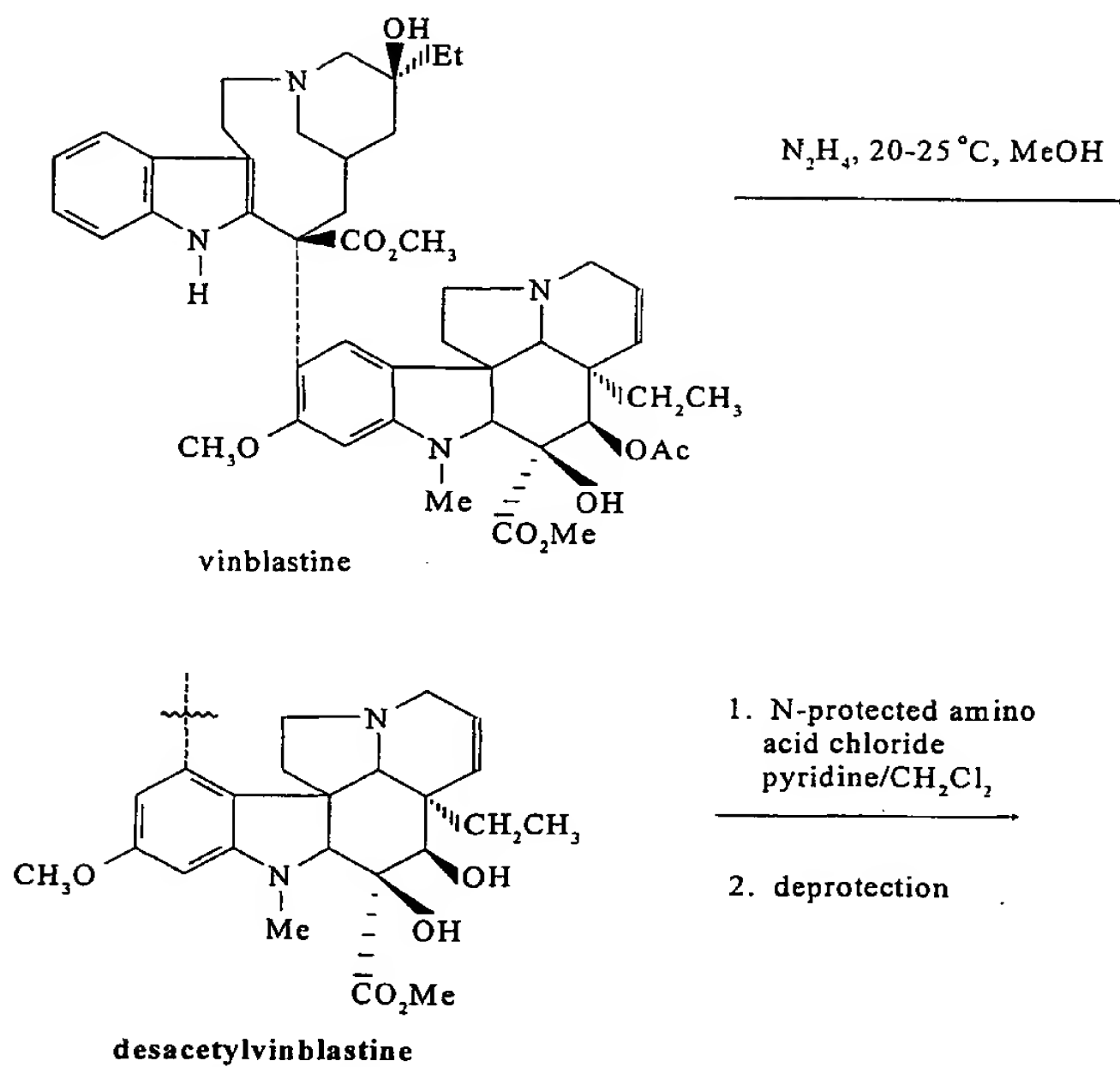
-153-

With respect to the embodiment of a peptidic substrate combined with desacetylvinblastine, the following Reaction Schemes IX and X illustrate the synthesis of the conjugates provided herein.

- Reaction Scheme IX illustrates preparation of conjugates provided
- 5    herein containing the peptidic substrates provided herein and the vinca alkaloid cytotoxic agent vinblastine wherein the attachment of the oxygen of the 4-desacetylvinblastine is at the C-terminus of the peptidic substrate. While other sequences of reactions can be useful in forming such conjugates, it is known that initial attachment of a single amino acid
- 10   to the 4-oxygen and subsequent attachment of the remaining peptidic substrate sequence to that amino acid is an exemplary method (see, International Patent Application Publication No. WO 99/28345). It also is known that 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine (ODHBT) can be utilized in place of HOAt in the final coupling step.
- 15        Reaction Scheme X illustrates preparation of conjugates of the peptidic substrates provided herein wherein a hydroxy alkanoyl acid is used as a linker between the vinca drug and the peptidic substrate.

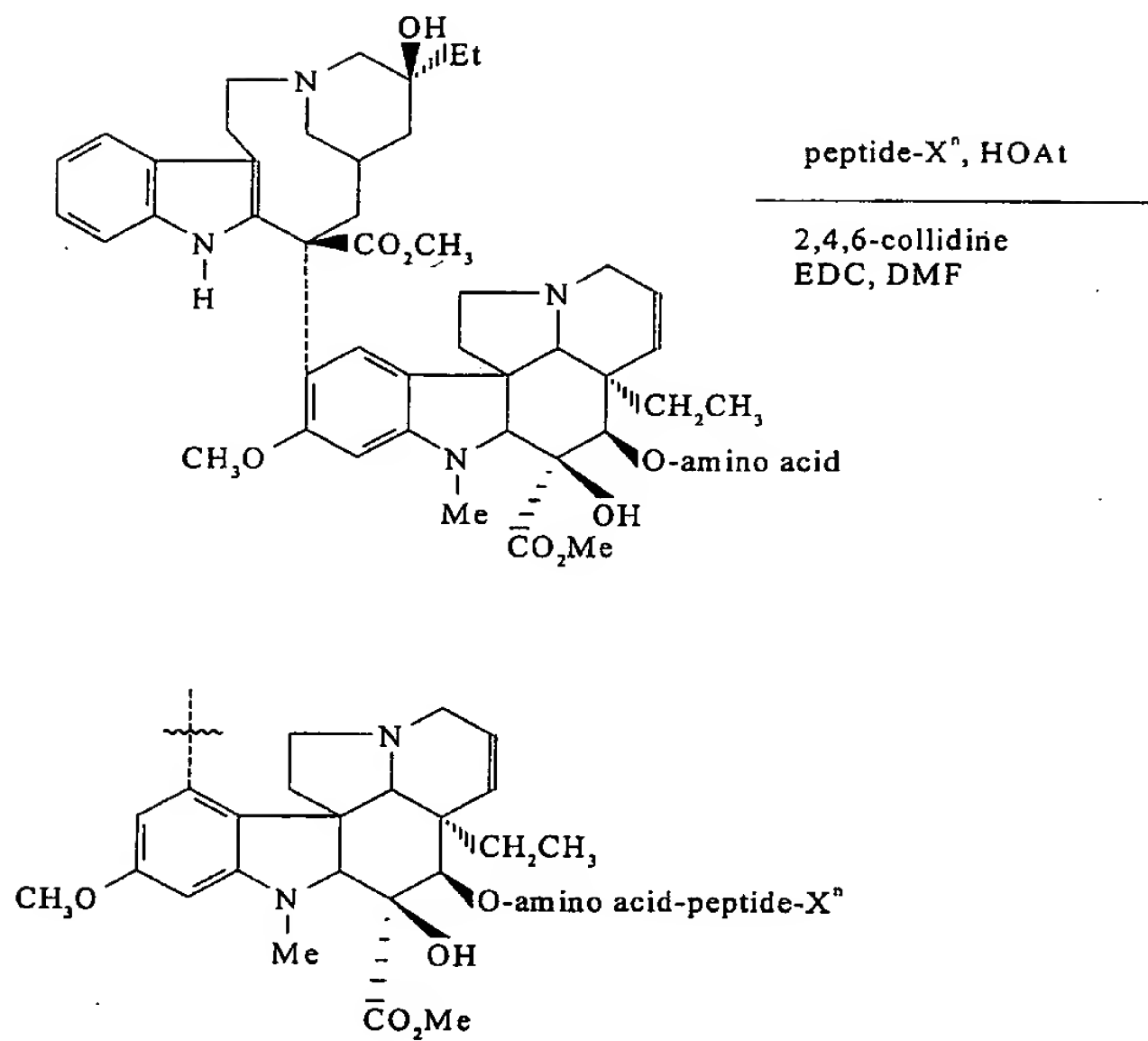
-154-

## REACTION SCHEME IX



-155-

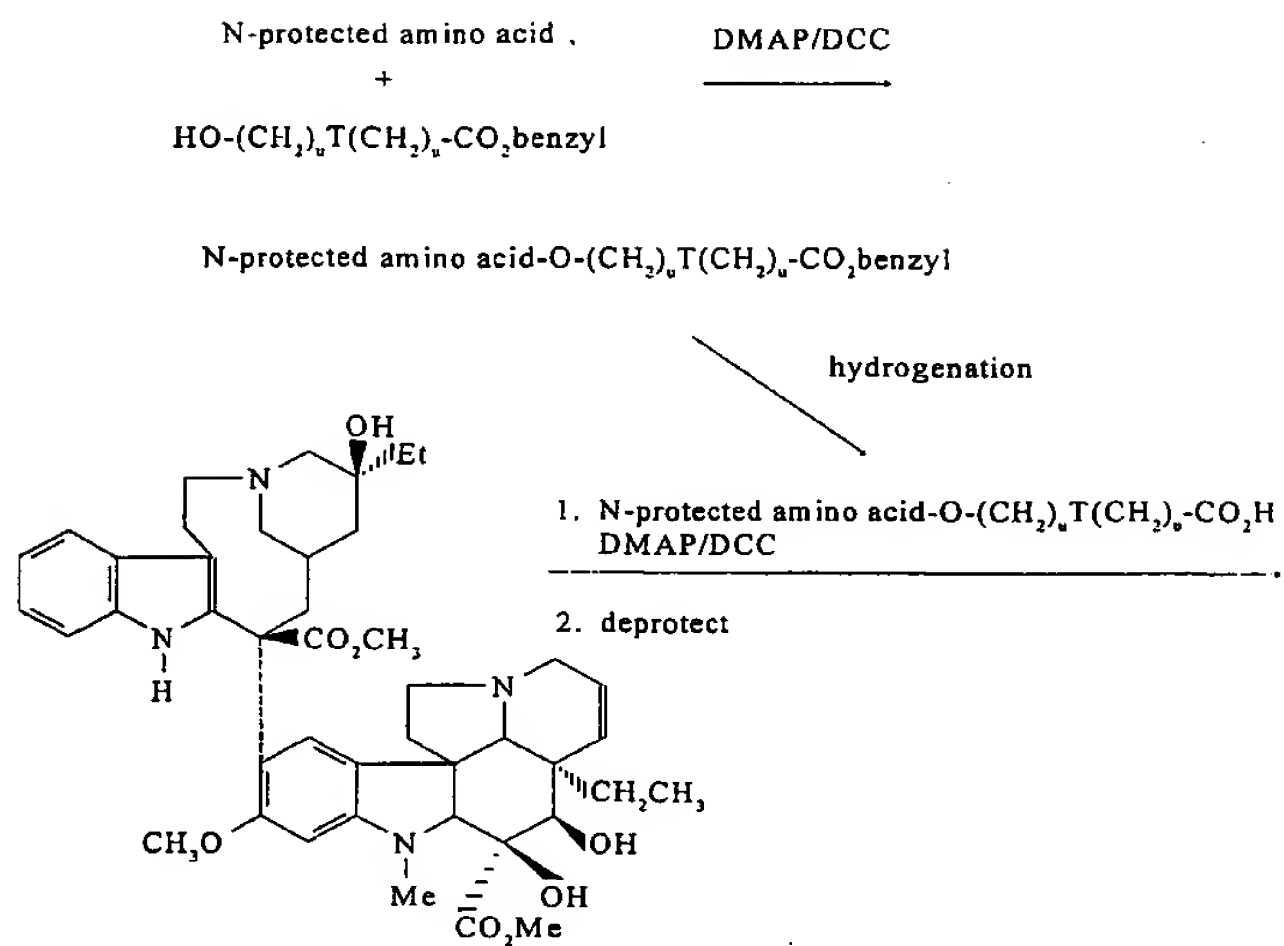
## REACTION SCHEME IX (Continued)





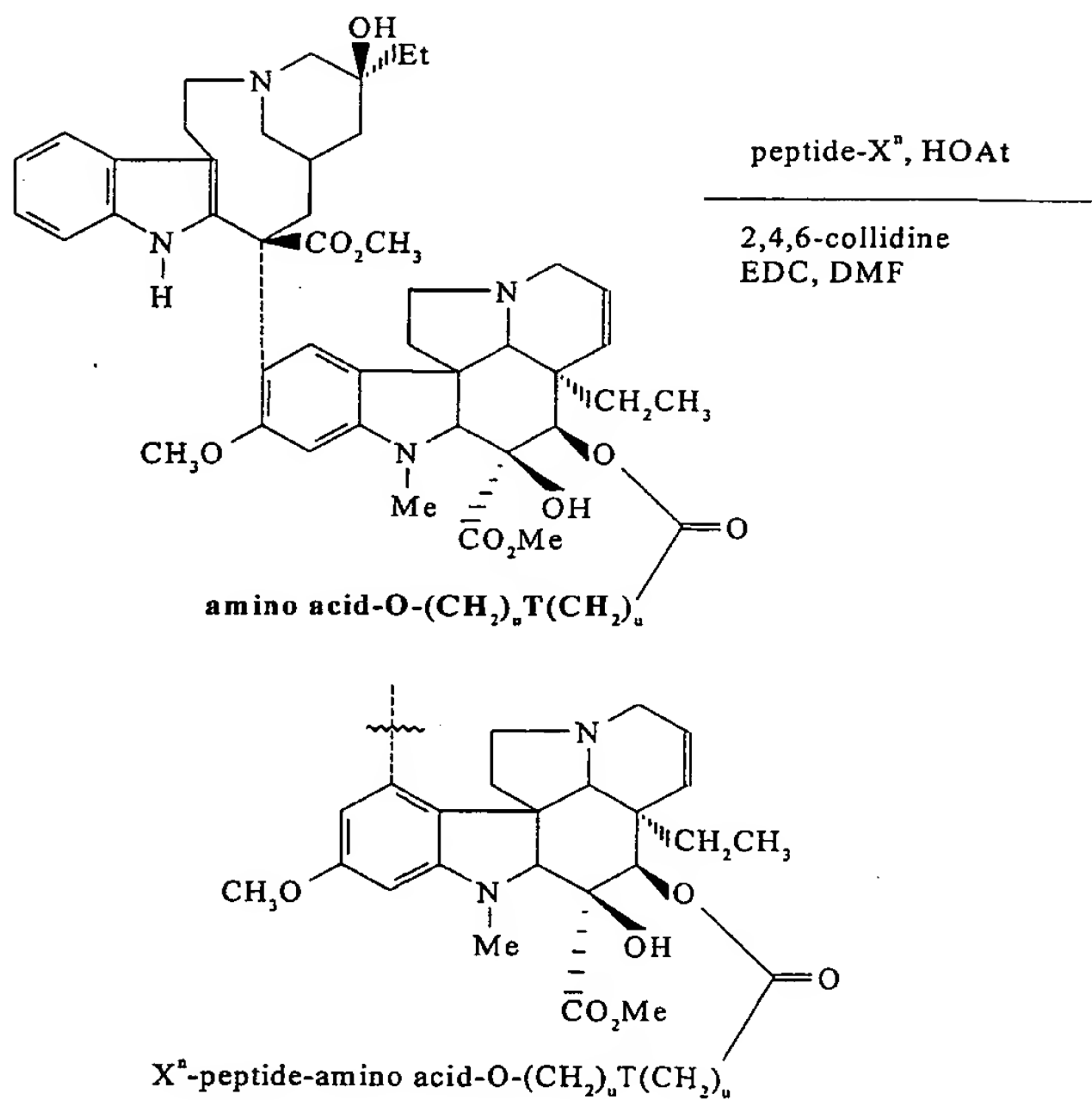
-156-

## REACTION SCHEME X



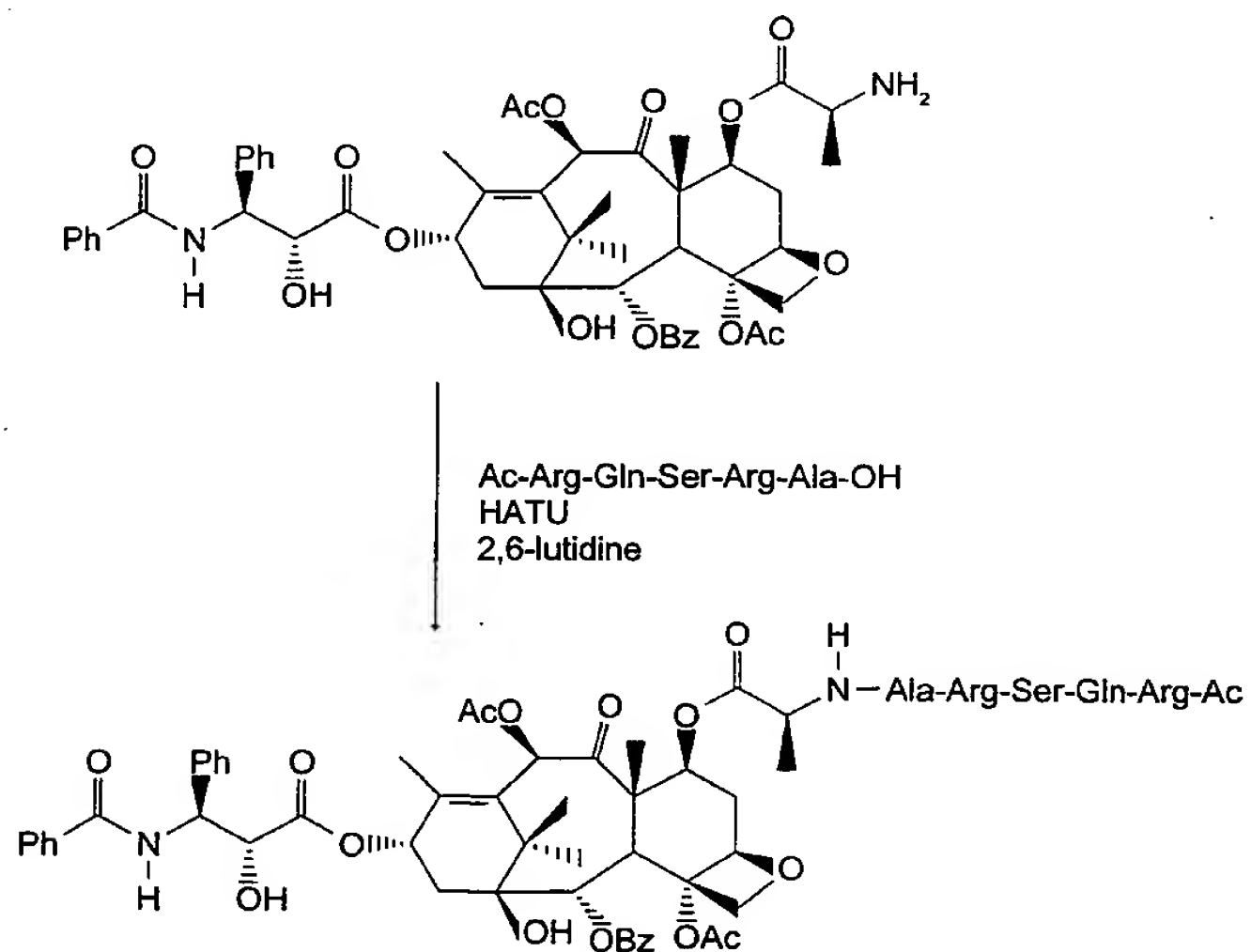
-157-

## REACTION SCHEME X (Continued)



-158-

Taxol conjugates provided herein may be prepared by the general method provided below. The preparation of 7-Ala-Taxol and 7-Gly-Taxol is disclosed in Mathew *et al.* (1992) *J. Med. Chem.* 35:145-151.



-159-

**E. Formulation and administration of pharmaceutical compositions**

The conjugates and compositions provided herein are used for treating, preventing, or ameliorating one or more symptoms of any disease or disorder that can be treated by targeting a cell or tissue that expresses a cell surface protease, particularly, a serine protease, on its surface at higher levels compared to other cells, or soluble, shed or released forms thereof. These include, but are not limited to, hyperproliferative diseases, such as cancer, any disease associated with aberrant or excessive angiogenesis, autoimmune disorders, inflammatory diseases and any other disease for which an appropriate cell surface protease, including cell-associated and cell-localized proteases, can be identified.

The pharmaceutical compositions provided herein contain therapeutically effective amounts of one or more of the conjugates provided herein that are useful in the prevention, treatment, or amelioration of one or more of the symptoms of diseases or disorders associated with undesired and/or uncontrolled angiogenesis or neovascularization. Such diseases or disorders include, but are not limited to, solid neoplasms, including lung, colon, esophageal, breast, ovarian and prostate cancers; vascular malformations and cardiovascular disorders, including, but not limited to, angiofibroma, angiolioma, atherosclerosis, restenosis/reperfusion injury, arteriovenous malformations, hemangiomatosis and vascular adhesions, dyschondroplasia with vascular hematomas, hereditary hemorrhagic telangiectasia and Von Hippel Lindau syndrome; chronic inflammatory diseases and adherent wound repairs, including, but not limited to, diabetes mellitus, hemophiliac joints, inflammatory bowel disease, nonhealing fractures, rapidly progressing periodontitis, juvenile periodontitis, psoriasis, rheumatoid arthritis, venous stasis ulcers,

-160-

- granulations-burns, hypertrophic scars, liver cirrhosis, osteoradionecrosis, postoperative adhesions, pyogenic granuloma and systemic sclerosis; circulatory disorders, including, but not limited to, Raynaud's phenomenon; crest syndromes, including, but not limited to, calcinosis,
- 5 esophageal, dyomotility, sclerodactyly and teangiectasis; dermatological disorders, including, but not limited to, systemic vasculitis, scleroderma, pyoderma gangrenosum, vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome and Osler-Weber-Rendu syndrome;
- 10 and ocular disorders, including, but not limited to, blindness caused by ocular neovascular disease, corneal graft neovascularization, macular degeneration in the eye, neovascular glaucoma, trachoma, diabetic retinopathy, myopic degeneration, retinopathy of prematurity, retrolental fibroplasia and corneal neovascularization.
- 15 The compositions contain one or more conjugates provided herein. The conjugates can be formulated into suitable pharmaceutical preparations such as, for example, solutions, suspensions, tablets, dispersible tablets, pills, capsules, powders, sustained release formulations or elixirs, for oral administration or in sterile solutions or
- 20 suspensions for parenteral administration, as well as transdermal patch preparation and dry powder inhalers. Typically the conjugates described above are formulated into pharmaceutical compositions using techniques and procedures well known in the art (see, *e.g.*, Ansel (1985) *Introduction to Pharmaceutical Dosage Forms*, Fourth Edition, p. 126)).
- 25 Effective concentrations can be empirically determined using animal models, *in vitro* models or test subjects.

In the compositions, effective concentrations of one or more conjugates or pharmaceutically acceptable derivatives thereof is (are) mixed with a suitable pharmaceutical carrier or vehicle. The conjugates

-161-

can be derivatized as the corresponding salts, esters, enol ethers or esters, acids, bases, solvates or hydrates prior to formulation, as described above. The concentrations of the conjugates in the compositions are effective for delivery of an amount, upon administration, 5 that treats, prevents, or ameliorates one or more of the symptoms of diseases or disorders associated with undesired and/or uncontrolled angiogenesis or neovascularization. Such diseases or disorders include, but are not limited to, solid neoplasms; vascular malformations and cardiovascular disorders, including, but not limited to, angiofibroma, 10 angiolipoma, atherosclerosis, restenosis/reperfusion injury, arteriovenous malformations, hemangiomatosis and vascular adhesions, dyschondroplasia with vascular hamartomas, hereditary hemorrhagic telangiectasia and Von Hippel Lindau syndrome; chronic inflammatory diseases and adherent wound repairs, including, but not limited to, 15 diabetes mellitus, hemophiliac joints, inflammatory bowel disease, nonhealing fractures, rapidly progressing periodontitis, juvenile periodontitis, psoriasis, rheumatoid arthritis, venous stasis ulcers, granulations-burns, hypertrophic scars, liver cirrhosis, osteoradionecrosis, postoperative adhesions, pyogenic granuloma and systemic sclerosis; 20 circulatory disorders, including, but not limited to, Raynaud's phenomenon; crest syndromes, including, but not limited to, calcinosis, esophageal, dyomotility, sclerodactyly and teangiectasis; dermatological disorders, including, but not limited to, systemic vasculitis, scleroderma, pyoderma gangrenosum, vasculopathy, venous, arterial ulcers, Sturge- 25 Weber syndrome, Port-wine stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome and Osler-Weber-Rendu syndrome; and ocular disorders, including, but not limited to, blindness caused by ocular neovascular disease, corneal graft neovascularization, macular degeneration in the eye, neovascular glaucoma, trachoma, diabetic

-162-

retinopathy, myopic degeneration, retinopathy of prematurity, retrolental fibroplasia and corneal neovascularization.

The conjugates herein can be formulated into pharmaceutical compositions suitable for topical, local, intravenous and systemic application. Effective concentrations of one or more of the conjugates are mixed with a suitable pharmaceutical carrier or vehicle. The concentrations or amounts of the conjugates that are effective requires delivery of an amount, upon administration, that ameliorates the symptoms or treats the disease. Typically, the compositions are formulated for single dosage administration. Therapeutically effective concentrations and amounts can be determined empirically by testing the conjugates in known *in vitro* and *in vivo* systems, such as those described here; dosages for humans or other animals can then be extrapolated therefrom.

Upon mixing or addition of the conjugate(s) with the vehicle, the resulting mixture can be a solution, suspension, emulsion or other such composition. The form of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the conjugate in the selected carrier or vehicle. The effective concentration is sufficient for ameliorating the symptoms of the disease, disorder or condition treated and can be empirically determined based upon *in vitro* and/or *in vivo* data, such as the data from the mouse xenograft model for tumors or rabbit ophthalmic model. If necessary, pharmaceutically acceptable salts or other derivatives of the conjugates can be prepared.

Pharmaceutical carriers or vehicles suitable for administration of the conjugates provided herein include any such carriers known to those skilled in the art to be suitable for the particular mode of administration.

-163-

In addition, the conjugates can be formulated as the sole pharmaceutically active ingredient in the composition or can be combined with other active ingredients.

The conjugates can be administered by any appropriate route, for example, orally, parenterally, intravenously, intradermally, subcutaneously, or topically, in liquid, semi-liquid or solid form and are formulated in a manner suitable for each route of administration. Exemplary modes of administration depend upon the indication treated. Dermatological and ophthalmologic indications will typically be treated locally; whereas, tumors and vascular proliferative disorders, will typically be treated by systemic, intradermal or intramuscular, modes of administration.

The conjugate is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. It is understood that number and degree of side effects depends upon the condition for which the conjugates are administered. For example, certain toxic and undesirable side effects are tolerated when treating life-threatening illnesses, such as tumors, that would not be tolerated when treating disorders of lesser consequence.

The concentration of conjugate in the composition will depend on absorption, inactivation and excretion rates thereof, the dosage schedule, and amount administered as well as other factors known to those of skill in the art.

Typically a therapeutically effective dosage should produce a serum concentration of active ingredient of from about 0.1 ng/ml to about 50-100  $\mu$ g/ml. The pharmaceutical compositions typically should provide a dosage of from about 0.01 mg to about 100 - 2000 mg of conjugate, depending upon the conjugate selected as adjusted for body surface area



-164-

and/or weight. Typically, for intravenous or systemic treatment a daily dosage of about between 0.05 and 0.5 mg/kg should be sufficient. Local application for ophthalmic disorders should provide about 1 ng up to 100  $\mu$ g, generally about 1  $\mu$ g to about 10  $\mu$ g, per single dosage

- 5 administration. It is understood that the amount to administer is a function of the conjugate selected, the indication treated, and possibly the side effects that will be tolerated. Dosages can be empirically determined using recognized models for each disorder.

- Typically, the compositions are formulated for single dosage
- 10 administration. To formulate a composition, the weight fraction of conjugate is dissolved, suspended, dispersed or otherwise mixed in a selected vehicle at an effective concentration such that the treated condition is relieved or ameliorated. Pharmaceutical carriers or vehicles suitable for administration of the conjugates provided herein include any
- 15 such carriers known to those skilled in the art to be suitable for the particular mode of administration.

- In addition, the conjugates can be formulated as the sole ingredient in the composition or can be combined with other active ingredients. Liposomal suspensions, including tissue-targeted liposomes, particularly
- 20 tumor-targeted liposomes, also can be suitable as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art. For example, liposome formulations can be prepared as described in U.S. Patent No. 4,522,811. Briefly, liposomes such as multilamellar vesicles (MLV's) can be formed by drying down egg
- 25 phosphatidyl choline and brain phosphatidyl serine (7:3 molar ratio) on the inside of a flask. A solution of a conjugate provided herein in phosphate buffered saline lacking divalent cations (PBS) is added and the flask shaken until the lipid film is dispersed. The resulting vesicles are

-165-

washed to remove unencapsulated conjugate, pelleted by centrifugation, and then resuspended in PBS.

The conjugate is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the  
5 absence of undesirable side effects on the patient treated. The therapeutically effective concentration can be determined empirically by testing the conjugates in *in vitro* and *in vivo* systems described herein (see, *e.g.*, EXAMPLES 3 and 4) and then extrapolated therefrom for dosages for humans.

10 The concentration of conjugate in the pharmaceutical composition will depend on absorption, inactivation and excretion rates of the conjugate, the physicochemical characteristics of the conjugate, the dosage schedule, and amount administered as well as other factors known to those of skill in the art. For example, the amount that is  
15 delivered is sufficient to ameliorate one or more of the symptoms of diseases or disorders associated with undesired and/or uncontrolled angiogenesis or neovascularization, as described herein.

Typically a therapeutically effective dosage should produce a serum concentration of active ingredient of from about 0.1 ng/ml to about 50-  
20 100 µg/ml. The pharmaceutical compositions typically should provide a dosage of from about 0.001 mg to about 2000 mg of conjugate per kilogram of body weight per day. Pharmaceutical dosage unit forms are prepared to provide from about 1 mg to about 1000 mg and generally from about 10 to about 500 mg of the essential active ingredient or a  
25 combination of essential ingredients per dosage unit form.

The conjugate can be administered at once, or can be divided into a number of smaller doses to be administered at intervals of time. It is understood that the precise dosage and duration of treatment is a function of the disease being treated and can be determined empirically

-166-

using known testing protocols or by extrapolation from *in vivo* or *in vitro* test data. It is to be noted that concentrations and dosage values can also vary with the severity of the condition to be alleviated. It is to be further understood that for any particular subject, specific dosage

5 regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the compositions, and that the concentration ranges set forth herein are exemplary only and are not intended to limit the scope or practice of the claimed compositions.

10 Exemplary pharmaceutically acceptable derivatives include acids, bases, enol ethers and esters, salts, esters, hydrates, solvates and conjugate forms. The derivative is selected such that its pharmacokinetic properties are superior to the corresponding neutral conjugate.

Thus, effective concentrations or amounts of one or more of the  
15 conjugates described herein or pharmaceutically acceptable derivatives thereof are mixed with a suitable pharmaceutical carrier or vehicle for systemic, topical or local administration to form pharmaceutical compositions. Conjugates are included in an amount effective for ameliorating one or more symptoms of, or for treating or preventing  
20 diseases or disorders associated with undesired and/or uncontrolled angiogenesis or neovascularization, as described herein. The concentration of conjugate in the composition will depend on absorption, inactivation, excretion rates of the conjugate, the dosage schedule, amount administered, particular formulation as well as other factors  
25 known to those of skill in the art.

The compositions are intended to be administered by a suitable route, including orally, parenterally, rectally, topically and locally. For oral administration, capsules and tablets are generally employed. The compositions are in liquid, semi-liquid or solid form and are formulated in

-167-

a manner suitable for each route of administration. Exemplary modes of administration include parenteral and oral modes of administration.

Solutions or suspensions used for parenteral, intradermal, subcutaneous, or topical application can include any of the following

5 components: a sterile diluent, such as water for injection, saline solution, fixed oil, polyethylene glycol, glycerine, propylene glycol or other synthetic solvent; antimicrobial agents, such as benzyl alcohol and methyl parabens; antioxidants, such as ascorbic acid and sodium bisulfite; chelating agents, such as ethylenediaminetetraacetic acid (EDTA);

10 buffers, such as acetates, citrates and phosphates; and agents for the adjustment of tonicity such as sodium chloride or dextrose. Parenteral preparations can be enclosed in ampules, disposable syringes or single or multiple dose vials made of glass, plastic or other suitable material.

In instances in which the conjugates exhibit insufficient solubility,

15 methods for solubilizing conjugates can be used. Such methods are known to those of skill in this art, and include, but are not limited to, using cosolvents, such as dimethylsulfoxide (DMSO), using surfactants, such as TWEEN®, or dissolution in aqueous sodium bicarbonate. Derivatives of the conjugates also can be used in formulating effective

20 pharmaceutical compositions.

Upon mixing or addition of the conjugate(s), the resulting mixture can be a solution, suspension, emulsion or the like. The form of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the conjugate in the

25 selected carrier or vehicle. The effective concentration is sufficient for ameliorating the symptoms of the disease, disorder or condition treated and can be empirically determined.

The pharmaceutical compositions are provided for administration to humans and animals in unit dosage forms, such as tablets, capsules, pills,

-168-

powders, granules, sterile parenteral solutions or suspensions, and oral solutions or suspensions, and oil-water emulsions containing suitable quantities of the conjugates or pharmaceutically acceptable derivatives thereof. The conjugates and derivatives thereof are typically formulated

5 and administered in unit-dosage forms or multiple-dosage forms.

Unit-dose forms as used herein refers to physically discrete units suitable for human and animal subjects and packaged individually as is known in the art. Each unit-dose contains a predetermined quantity of the conjugate sufficient to produce the desired therapeutic effect, in

10 association with the required pharmaceutical carrier, vehicle or diluent. Examples of unit-dose forms include ampoules and syringes and individually packaged tablets or capsules. Unit-dose forms can be administered in fractions or multiples thereof. A multiple-dose form is a plurality of identical unit-dosage forms packaged in a single container to

15 be administered in segregated unit-dose form. Examples of multiple-dose forms include vials, bottles of tablets or capsules or bottles of pints or gallons. Hence, multiple dose form is a multiple of unit-doses which are not segregated in packaging.

The composition can contain along with the conjugate: a diluent

20 such as lactose, sucrose, dicalcium phosphate, or carboxymethyl-cellulose; a lubricant, such as magnesium stearate, calcium stearate and talc; and a binder such as starch, natural gums, such as gum acacia-gelatin, glucose, molasses, polyvinylpyrrolidone, celluloses and derivatives thereof, povidone, crospovidones and other such binders known to those

25 of skill in the art. Liquid pharmaceutically administrable compositions can, for example, be prepared by dissolving, dispersing, or otherwise mixing a conjugate as defined above and optional pharmaceutical adjuvants in a carrier, such as, for example, water, saline, aqueous dextrose, glycerol, glycols, ethanol, and the like, to thereby form a

-169-

- solution or suspension. If desired, the pharmaceutical composition to be administered can also contain minor amounts of nontoxic auxiliary substances such as wetting agents, emulsifying agents, or solubilizing agents, pH buffering agents and the like, for example, acetate, sodium
- 5 citrate, cyclodextrine derivatives, sorbitan monolaurate, triethanolamine sodium acetate, triethanolamine oleate, and other such agents. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in this art; for example, see Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., 15th Edition, 1975.
- 10 The composition or formulation to be administered will, in any event, contain a quantity of the conjugate in an amount sufficient to alleviate the symptoms of the treated subject.

- Dosage forms or compositions containing active ingredient in the range of 0.005% to 100% with the balance made up from non-toxic
- 15 carrier can be prepared. For oral administration, a pharmaceutically acceptable non-toxic composition is formed by the incorporation of any of the normally employed excipients, such as, for example pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, talcum, cellulose derivatives, sodium crosscarmellose, glucose, sucrose,
- 20 magnesium carbonate or sodium saccharin. Such compositions include solutions, suspensions, tablets, capsules, powders and sustained release formulations, such as, but not limited to, implants and microencapsulated delivery systems, and biodegradable, biocompatible polymers, such as collagen, ethylene vinyl acetate, polyanhydrides, polyglycolic acid,
- 25 polyorthoesters, polylactic acid and others. Methods for preparation of these compositions are known to those skilled in the art. The contemplated compositions can contain 0.001%-100% active ingredient, such as 0.1-85%, for example 75-95%.

-170-

The conjugates or pharmaceutically acceptable derivatives can be prepared with carriers that protect the conjugate against rapid elimination from the body, such as time release formulations or coatings. The compositions can include other conjugates to obtain desired combinations of properties. The conjugates provided herein, or pharmaceutically acceptable derivatives thereof as described herein, also can be advantageously administered for therapeutic or prophylactic purposes together with another pharmacological agent known in the general art to be of value in treating one or more of the diseases or medical conditions referred to hereinabove, such as diseases or disorders associated with undesired and/or uncontrolled angiogenesis or neovascularization. It is to be understood that such combination therapy constitutes a further aspect of the compositions and methods of treatment provided herein.

#### 1. Compositions for oral administration

Oral pharmaceutical dosage forms are either solid, gel or liquid. The solid dosage forms are tablets, capsules, granules, and bulk powders. Types of oral tablets include compressed, chewable lozenges and tablets which can be enteric-coated, sugar-coated or film-coated. Capsules can be hard or soft gelatin capsules, while granules and powders can be provided in non-effervescent or effervescent form with the combination of other ingredients known to those skilled in the art.

In certain embodiments, the formulations are solid dosage forms, such as, for example, capsules or tablets. The tablets, pills, capsules, troches and other dosage forms can contain, for example, any of the following ingredients, or compounds of a similar nature: a binder; a diluent; a disintegrating agent; a lubricant; a glidant; a sweetening agent; and a flavoring agent.

Examples of binders include microcrystalline cellulose, gum tragacanth, glucose solution, acacia mucilage, gelatin solution, sucrose

-171-

and starch paste. Lubricants include talc, starch, magnesium or calcium stearate, lycopodium and stearic acid. Diluents include, for example, lactose, sucrose, starch, kaolin, salt, mannitol and dicalcium phosphate. Glidants include, but are not limited to, colloidal silicon dioxide.

- 5 Disintegrating agents include crosscarmellose sodium, sodium starch glycolate, alginic acid, corn starch, potato starch, bentonite, methylcellulose, agar and carboxymethylcellulose. Coloring agents include, for example, any of the approved certified water soluble FD and C dyes, mixtures thereof; and water insoluble FD and C dyes suspended
- 10 on alumina hydrate. Sweetening agents include sucrose, lactose, mannitol and artificial sweetening agents such as saccharin, and any number of spray dried flavors. Flavoring agents include natural flavors extracted from plants such as fruits and synthetic blends of conjugates which produce a pleasant sensation, such as, but not limited to peppermint and
- 15 methyl salicylate. Wetting agents include propylene glycol monostearate, sorbitan monooleate, diethylene glycol monolaurate and polyoxyethylene laural ether. Emetic-coatings include fatty acids, fats, waxes, shellac, ammoniated shellac and cellulose acetate phthalates. Film coatings include hydroxyethylcellulose, sodium carboxymethylcellulose,
- 20 polyethylene glycol 4000 and cellulose acetate phthalate.

If oral administration is desired, the conjugate could be provided in a composition that protects it from the acidic environment of the stomach. For example, the composition can be formulated in an enteric coating that maintains its integrity in the stomach and releases the

25 conjugate in the intestine. The composition also can be formulated in combination with an antacid or other such ingredient.

When the dosage unit form is a capsule, it can contain, in addition to material of the above type, a liquid carrier such as a fatty oil. In addition, dosage unit forms can contain various other materials which



-172-

modify the physical form of the dosage unit, for example, coatings of sugar and other enteric agents. The conjugates also can be administered as a component of an elixir, suspension, syrup, wafer, sprinkle, chewing gum or the like. A syrup can contain, in addition to the conjugates,  
5 sucrose as a sweetening agent and certain preservatives, dyes and colorings and flavors.

The conjugates also can be mixed with other active materials which do not impair the desired action, or with materials that supplement the desired action, such as antacids, H<sub>2</sub> blockers, and diuretics. Higher  
10 concentrations, up to about 98% by weight of the conjugate can be included.

Pharmaceutically acceptable carriers included in tablets are binders, lubricants, diluents, disintegrating agents, coloring agents, flavoring agents, and wetting agents. Enteric-coated tablets, because of the  
15 enteric-coating, resist the action of stomach acid and dissolve or disintegrate in the neutral or alkaline intestines. Sugar-coated tablets are compressed tablets to which different layers of pharmaceutically acceptable substances are applied. Film-coated tablets are compressed tablets which have been coated with a polymer or other  
20 suitable coating. Multiple compressed tablets are compressed tablets made by more than one compression cycle utilizing the pharmaceutically acceptable substances previously mentioned. Coloring agents also can be used in the above dosage forms. Flavoring and sweetening agents are used in compressed tablets, sugar-coated, multiple compressed and  
25 chewable tablets. Flavoring and sweetening agents are especially useful in the formation of chewable tablets and lozenges.

Liquid oral dosage forms include aqueous solutions, emulsions, suspensions, solutions and/or suspensions reconstituted from

-173-

non-effervescent granules and effervescent preparations reconstituted from effervescent granules. Aqueous solutions include, for example, elixirs and syrups. Emulsions are either oil-in-water or water-in-oil.

Elixirs are clear, sweetened, hydroalcoholic preparations.

- 5 Pharmaceutically acceptable carriers used in elixirs include solvents. Syrups are concentrated aqueous solutions of a sugar, for example, sucrose, and can contain a preservative. An emulsion is a two-phase system in which one liquid is dispersed in the form of small globules throughout another liquid. Pharmaceutically acceptable carriers used in
- 10 emulsions are non-aqueous liquids, emulsifying agents and preservatives. Suspensions use pharmaceutically acceptable suspending agents and preservatives. Pharmaceutically acceptable substances used in non-effervescent granules, to be reconstituted into a liquid oral dosage form, include diluents, sweeteners and wetting agents. Pharmaceutically
- 15 acceptable substances used in effervescent granules, to be reconstituted into a liquid oral dosage form, include organic acids and a source of carbon dioxide. Coloring and flavoring agents are used in all of the above dosage forms.

Solvents include glycerin, sorbitol, ethyl alcohol and syrup.

- 20 Examples of preservatives include glycerin, methyl and propylparaben, benzoic acid, sodium benzoate and alcohol. Examples of non-aqueous liquids utilized in emulsions include mineral oil and cottonseed oil. Examples of emulsifying agents include gelatin, acacia, tragacanth, bentonite, and surfactants such as polyoxyethylene sorbitan monooleate.
- 25 Suspending agents include sodium carboxymethylcellulose, pectin, tragacanth, Veegum and acacia. Diluents include lactose and sucrose. Sweetening agents include sucrose, syrups, glycerin and artificial sweetening agents such as saccharin. Wetting agents include propylene glycol monostearate, sorbitan monooleate, diethylene glycol monolaurate

-174-

and polyoxyethylene lauryl ether. Organic adds include citric and tartaric acid. Sources of carbon dioxide include sodium bicarbonate and sodium carbonate. Coloring agents include any of the approved certified water soluble FD and C dyes, and mixtures thereof. Flavoring agents include  
5 natural flavors extracted from plants such fruits, and synthetic blends of conjugates which produce a pleasant taste sensation.

For a solid dosage form, the solution or suspension, in for example propylene carbonate, vegetable oils or triglycerides, for example the formulation can be encapsulated in a gelatin capsule. Such solutions, and  
10 the preparation and encapsulation thereof, are disclosed in U.S. Patent Nos 4,328,245; 4,409,239; and 4,410,545. For a liquid dosage form, the solution, *e.g.*, for example, in a polyethylene glycol, can be diluted with a sufficient quantity of a pharmaceutically acceptable liquid carrier, *e.g.*, water, to be easily measured for administration.

15 Alternatively, liquid or semi-solid oral formulations can be prepared by dissolving or dispersing the conjugate or derivative thereof in vegetable oils, glycols, triglycerides, propylene glycol esters (*e.g.*, propylene carbonate) and other such carriers, and encapsulating these solutions or suspensions in hard or soft gelatin capsule shells. Other  
20 useful formulations include those set forth in U.S. Patent Nos. Re 28,819 and 4,358,603. Briefly, such formulations include, but are not limited to, those containing a conjugate provided herein, a dialkylated mono- or poly-alkylene glycol, including, but not limited to, 1,2-dimethoxymethane, diglyme, triglyme, tetraglyme, polyethylene glycol-350-dimethyl ether,  
25 polyethylene glycol-550-dimethyl ether, polyethylene glycol-750-dimethyl ether wherein 350, 550 and 750 refer to the approximate average molecular weight of the polyethylene glycol, and one or more antioxidants, such as butylated hydroxytoluene (BHT), butylated hydroxyanisole (BHA), propyl gallate, vitamin E, hydroquinone,

-175-

hydroxycoumarins, ethanolamine, lecithin, cephalin, ascorbic acid, malic acid, sorbitol, phosphoric acid, thiodipropionic acid and its esters, and dithiocarbamates.

- Other formulations include, but are not limited to, aqueous
- 5 alcoholic solutions including a pharmaceutically acceptable acetal. Alcohols used in these formulations are any pharmaceutically acceptable water-miscible solvents having one or more hydroxyl groups, including, but not limited to, propylene glycol and ethanol. Acetals include, but are not limited to, di(lower alkyl) acetals of lower alkyl aldehydes such as
- 10 acetaldehyde diethyl acetal.

- In all embodiments, tablets and capsules formulations can be coated as known by those of skill in the art in order to modify or sustain dissolution of the conjugate. Thus, for example, they can be coated with a conventional enterically digestible coating, such as phenylsalicylate,
- 15 waxes and cellulose acetate phthalate.

## **2. Injectables, solutions and emulsions**

- Parenteral administration, generally characterized by injection, either subcutaneously, intramuscularly or intravenously also is contemplated herein. Injectables can be prepared in conventional forms,
- 20 either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as emulsions. Suitable excipients are, for example, water, saline, dextrose, glycerol or ethanol. In addition, if desired, the pharmaceutical compositions to be administered can also contain minor amounts of non-toxic auxiliary
- 25 substances such as wetting or emulsifying agents, pH buffering agents, stabilizers, solubility enhancers, and other such agents, such as for example, sodium acetate, sorbitan monolaurate, triethanolamine oleate and cyclodextrins. Implantation of a slow-release or sustained-release system, such that a constant level of dosage is maintained (see, *e.g.*,

-176-

U.S. Patent No. 3,710,795) also is contemplated herein. Briefly, a conjugate provided herein is dispersed in a solid inner matrix, *e.g.*, polymethylmethacrylate, polybutylmethacrylate, plasticized or unplasticized polyvinylchloride, plasticized nylon, plasticized

5 polyethyleneterephthalate, natural rubber, polyisoprene, polyisobutylene, polybutadiene, polyethylene, ethylene-vinylacetate copolymers, silicone rubbers, polydimethylsiloxanes, silicone carbonate copolymers, hydrophilic polymers such as hydrogels of esters of acrylic and methacrylic acid, collagen, cross-linked polyvinylalcohol and cross-linked

10 partially hydrolyzed polyvinyl acetate, that is surrounded by an outer polymeric membrane, *e.g.*, polyethylene, polypropylene, ethylene/propylene copolymers, ethylene/ethyl acrylate copolymers, ethylene/vinylacetate copolymers, silicone rubbers, polydimethyl siloxanes, neoprene rubber, chlorinated polyethylene, polyvinylchloride,

15 vinylchloride copolymers with vinyl acetate, vinylidene chloride, ethylene and propylene, ionomer polyethylene terephthalate, butyl rubber epichlorohydrin rubbers, ethylene/vinyl alcohol copolymer, ethylene/vinyl acetate/vinyl alcohol terpolymer, and ethylene/vinyloxyethanol copolymer, that is insoluble in body fluids. The conjugate diffuses through the outer

20 polymeric membrane in a release rate controlling step. The percentage of conjugate contained in such parenteral compositions is highly dependent on the specific nature thereof, as well as the activity of the conjugate and the needs of the subject.

Parenteral administration of the compositions includes intravenous,

25 subcutaneous and intramuscular administrations. Preparations for parenteral administration include sterile solutions ready for injection, sterile dry soluble products, such as lyophilized powders, ready to be combined with a solvent just prior to use, including hypodermic tablets, sterile suspensions ready for injection, sterile dry insoluble products ready

-177-

to be combined with a vehicle just prior to use and sterile emulsions. The solutions can be either aqueous or nonaqueous.

If administered intravenously, suitable carriers include physiological saline or phosphate buffered saline (PBS), and solutions containing  
5 thickening and solubilizing agents, such as glucose, polyethylene glycol, and polypropylene glycol and mixtures thereof.

Pharmaceutically acceptable carriers used in parenteral preparations include aqueous vehicles, nonaqueous vehicles, antimicrobial agents, isotonic agents, buffers, antioxidants, local anesthetics, suspending and  
10 dispersing agents, emulsifying agents, sequestering or chelating agents and other pharmaceutically acceptable substances.

Examples of aqueous vehicles include Sodium Chloride Injection, Ringers Injection, Isotonic Dextrose Injection, Sterile Water Injection, Dextrose and Lactated Ringers Injection. Nonaqueous parenteral vehicles  
15 include fixed oils of vegetable origin, cottonseed oil, corn oil, sesame oil and peanut oil. Antimicrobial agents in bacteriostatic or fungistatic concentrations must be added to parenteral preparations packaged in multiple-dose containers which include phenols or cresols, mercurials, benzyl alcohol, chlorobutanol, methyl and propyl p-hydroxybenzoic acid  
20 esters, thimerosal, benzalkonium chloride and benzethonium chloride. Isotonic agents include sodium chloride and dextrose. Buffers include phosphate and citrate. Antioxidants include sodium bisulfate. Local anesthetics include procaine hydrochloride. Suspending and dispersing agents include sodium carboxymethylcellulose, hydroxypropyl  
25 methylcellulose and polyvinylpyrrolidone. Emulsifying agents include Polysorbate 80 (TWEEN® 80). A sequestering or chelating agent of metal ions include EDTA. Pharmaceutical carriers also include ethyl alcohol, polyethylene glycol and propylene glycol for water miscible vehicles and

-178-

sodium hydroxide, hydrochloric acid, citric acid or lactic acid for pH adjustment.

The concentration of the conjugate is adjusted so that an injection provides an effective amount to produce the desired pharmacological effect. The exact dose depends on the age, weight and condition of the patient or animal as is known in the art.

The unit-dose parenteral preparations are packaged in an ampule, a vial or a syringe with a needle. All preparations for parenteral administration must be sterile, as is known and practiced in the art.

10 Illustratively, intravenous or intraarterial infusion of a sterile aqueous solution containing a conjugate is an effective mode of administration. Another embodiment is a sterile aqueous or oily solution or suspension containing a conjugate injected as necessary to produce the desired pharmacological effect.

15 Injectables are designed for local and systemic administration. Typically a therapeutically effective dosage is formulated to contain a concentration of at least about 0.1% w/w up to about 90% w/w or more, generally more than 1% w/w of the conjugate to the treated tissue(s). The conjugate can be administered at once, or can be divided  
20 into a number of smaller doses to be administered at intervals of time. It is understood that the precise dosage and duration of treatment is a function of the tissue being treated and can be determined empirically using known testing protocols or by extrapolation from *in vivo* or *in vitro* test data. It is to be noted that concentrations and dosage values can  
25 also vary with the age of the individual treated. It is to be further understood that for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the formulations, and that the concentration ranges set

-179-

forth herein are exemplary only and are not intended to limit the scope or practice of the claimed formulations.

The conjugate can be suspended in micronized or other suitable form or can be derivatized to produce a more soluble product. The form  
5 of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the conjugate in the selected carrier or vehicle. The effective concentration is sufficient for ameliorating the symptoms of the condition and can be empirically determined.

10           **3. Lyophilized powders**

Of interest herein are also lyophilized powders, which can be reconstituted for administration as solutions, emulsions and other mixtures. They also can be reconstituted and formulated as solids or gels.

15           The sterile, lyophilized powder is prepared by dissolving a conjugate provided herein, or a pharmaceutically acceptable derivative thereof, in a suitable solvent. The solvent can contain an excipient which improves the stability or other pharmacological component of the powder or reconstituted solution, prepared from the powder. Excipients that can  
20 be used include, but are not limited to, dextrose, sorbital, fructose, corn syrup, xylitol, glycerin, glucose, sucrose or other suitable agent. The solvent can also contain a buffer, such as citrate, sodium or potassium phosphate or other such buffer known to those of skill in the art at, typically, about neutral pH. Subsequent sterile filtration of the solution  
25 followed by lyophilization under standard conditions known to those of skill in the art provides the desired formulation. Generally, the resulting solution will be apportioned into vials for lyophilization. Each vial will contain a single dosage (such as 10-1000 mg, for example 100-500 mg) or multiple dosages of the conjugate. The lyophilized powder can be



-180-

stored under appropriate conditions, such as at about 4 °C to room temperature.

Reconstitution of this lyophilized powder with water for injection provides a formulation for use in parenteral administration. For  
5 reconstitution, generally about 1-50 mg, such 5-35 mg or about 9-30 mg of lyophilized powder, is added per mL of sterile water or other suitable carrier. The precise amount depends upon the selected conjugate, intended subject, and other empirically determinable parameters. Hence the amount can be empirically determined.

#### 10 4. Topical administration

Topical mixtures are prepared as described for the local and systemic administration. The resulting mixture can be a solution, suspension, emulsions or the like and are formulated as creams, gels, ointments, emulsions, solutions, elixirs, lotions, suspensions, tinctures,  
15 pastes, foams, aerosols, irrigations, sprays, suppositories, bandages, dermal patches or any other formulations suitable for topical administration.

The conjugates or pharmaceutically acceptable derivatives thereof can be formulated as aerosols for topical application, such as by  
20 inhalation (see, *e.g.*, U.S. Patent Nos. 4,044,126, 4,414,209, and 4,364,923, which describe aerosols for delivery of a steroid useful for treatment of inflammatory diseases, particularly asthma). These formulations for administration to the respiratory tract can be in the form of an aerosol or solution for a nebulizer, or as a microfine powder for  
25 insufflation, alone or in combination with an inert carrier such as lactose. In such a case, the particles of the formulation will typically have diameters of less than 50 microns, generally less than 10 microns.

The conjugates can be formulated for local or topical application, such as for topical application to the skin and mucous membranes, such

-181-

as in the eye, in the form of gels, creams, and lotions and for application to the eye or for intracisternal or intraspinal application. Topical administration is contemplated for transdermal delivery and also for administration to the eyes or mucosa, or for inhalation therapies. Nasal solutions of the conjugate alone or in combination with other pharmaceutically acceptable excipients also can be administered.

These solutions, particularly those intended for ophthalmic use, can be formulated as 0.01% - 10% isotonic solutions, pH about 5-7, with appropriate salts.

**5. Compositions for other routes of administration**

Other routes of administration, such as topical application, transdermal patches, and rectal administration are also contemplated herein.

For example, pharmaceutical dosage forms for rectal administration are rectal suppositories, capsules and tablets for systemic effect. Rectal suppositories are used herein mean solid bodies for insertion into the rectum which melt or soften at body temperature releasing one or more conjugates. Pharmaceutically acceptable substances utilized in rectal suppositories are bases or vehicles and agents to raise the melting point. Examples of bases include cocoa butter (theobroma oil), glycerin-gelatin, carbowax (polyoxyethylene glycol) and appropriate mixtures of mono-, di- and triglycerides of fatty acids. Combinations of the various bases can be used. Agents to raise the melting point of suppositories include spermaceti and wax. Rectal suppositories can be prepared either by the compressed method or by molding. The typical weight of a rectal suppository is about 2 to 3 gm.

Tablets and capsules for rectal administration are manufactured using the same pharmaceutically acceptable substance and by the same methods as for formulations for oral administration.

## 6. Articles of manufacture

The conjugates or pharmaceutically acceptable derivatives can be packaged as articles of manufacture containing packaging material, a conjugate or pharmaceutically acceptable derivative thereof provided  
5 herein, which is used for treatment, prevention or amelioration of one or more symptoms associated with proliferative diseases or disorders, and a label that indicates that the conjugate or pharmaceutically acceptable derivative thereof is used for treatment, prevention or amelioration of one or more symptoms associated with proliferative diseases or disorders.

10 The articles of manufacture provided herein contain packaging materials. Packaging materials for use in packaging pharmaceutical products are well known to those of skill in the art. See, *e.g.*, U.S. Patent Nos. 5,323,907, 5,052,558 and 5,033,252. Examples of pharmaceutical packaging materials include, but are not limited to, blister  
15 packs, bottles, tubes, inhalers, pumps, bags, vials, containers, syringes, bottles, and any packaging material suitable for a selected formulation and intended mode of administration and treatment. A wide array of formulations of the conjugates and compositions provided herein are contemplated as are a variety of treatments for any disorder in which a  
20 cell surface protease, or a soluble, shed or secreted form thereof, is implicated.

## F. Evaluation of the activity of the conjugates

Standard physiological, pharmacological and biochemical procedures are available for testing the conjugates to identify those that  
25 possess therapeutic activity upon action of a cell surface protease or a soluble, shed, or released form thereof. *In vitro* and *in vivo* assays that can be used to evaluate therapeutic activity, such as cytotoxicity, of the conjugates will depend upon the therapeutic agent being tested.

-183-

Exemplary assays are discussed briefly below with reference to cytotoxic conjugates (see, also, Examples). It is understood that the particular activity assayed will depend upon the conjugated therapeutic agent.

5           1.     *In vitro* Assays

The therapeutic activity, such as cytotoxicity, of the conjugates provided herein can be assessed by any assays normally used for assessing the therapeutic activity, such as cytotoxicity, of the unconjugated therapeutic agent. Numerous such assays are known, for example,  
10   assays can employ cells that express the targeted cell surface protease and the therapeutic activity of the therapeutic agent is assessed. For example, cytotoxicity can be assessed by measuring cell viability or by measuring cell proliferation, such as by incorporation of a labeled nucleotide or other such label. Generally the activity is compared with  
15   cells that do not express the targeted protease.

For example, the cells will be any that express a targeted MTSP or endotheliase. Such cells can be obtained by choosing cells known to express the cell surface protease, such as by determining tissue expression profiles, as discussed above, or by screening a variety of cell  
20   lines with an antibody for a targeted protease, or for the protease activity in the presence of a labeled, such as a chromogenic, substrate for the protease in the presence and absence of a known inhibitor of the targeted protease.

Alternatively, nucleic acid encoding the protease can be introduced  
25   in a cell line that does not express the protease, and expressed therein to produce a cell line that expresses the protease of interest. The resulting recombinant cells can be used in cytotoxicity assays.

-184-

## 2. *In vivo* Assays

Numerous animal models for assessing therapeutic activity are known. Any suitable *in vivo* model can be used. Exemplary are the mouse xenograft model and chicken embryo models.

### 5                   Chicken Embryo Model

The CAM model (chick embryo chorioallantoic membrane model; Ossowski (1988) *J. Cell Biol.* 107:2437-2445), provides another method for evaluating the inhibitory activity of a test compound. In the CAM model, tumor cells invade through the chorioallantoic membrane  
10 containing CAM (with tumor cells in the presence of several serine protease inhibitors results in less or no invasion of the tumor cells through the membrane). Thus, the CAM assay is performed with CAM and tumor cells in the presence and absence of various concentrations of test compound. The invasiveness of tumor cells is measured under such  
15 conditions to provide an indication of the compound's inhibitory activity. A compound having inhibitory activity correlates with less tumor invasion.

Thus, the CAM assay is performed with CAM and tumor cells in the presence and absence of various concentrations of a test compound. A compound having activity correlates with a change in tumor invasion  
20 and/or tumor growth.

For example, the ability of a cell surface protease to liberate a therapeutic agent, such as a cytotoxic agent, or the activity of a conjugate agent can be assessed using this model. If the therapeutic agent is released from the compound and it is an inhibitory agent there  
25 will be less tumor invasion or a decrease in size of the tumor. If the therapeutic agent is inactive in the conjugate, there will be no effect on tumor invasion.

-185-

The CAM model also is used in a standard assay of angiogenesis (*i.e.*, effect on formation of new blood vessels (Brooks *et al.* (1991) *Methods in Molecular Biology* 129:257-269). According to this model, a filter disc containing an angiogenesis inducer, such as basic fibroblast growth factor (bFGF) is placed onto the CAM. Diffusion of the cytokine into the CAM induces local angiogenesis, which can be measured in several ways such as by counting the number of blood vessel branch points within the CAM directly below the filter disc. The ability of identified compounds to inhibit cytokine-induced angiogenesis can be tested using this model. A test compound can either be added to the filter disc that contains the angiogenesis inducer, be placed directly on the membrane or be administered systemically. The extent of new blood vessel formation in the presence and/or absence of test compound can be compared using this model. The formation of fewer new blood vessels in the presence of a test compound would be indicative of anti-angiogenesis activity.

This can be adapted for use with the conjugates herein to 1) assess the activity of a therapeutic agent in the conjugate; and 2) to assess the ability of a particular cell surface protease to liberate a therapeutic agent from a conjugate.

#### Mouse xenograft model

*In vivo* activity can be assessed using recognized animal models, such as the well-known mouse xenograft model for anti-tumor activity (see, *e.g.*, Beitz *et al.* (1992) *Cancer Research* 52:227-230; Houghton *et al.* (1982) *Cancer Res.* 42:535-539; Bogden *et al.* (1981) *Cancer (Philadelphia)* 48:10-20; Hoogenhout *et al.* (1983) *Int. J. Radiat. Oncol., Biol. Phys.* 9:871-879; Stastny *et al.* (1993) *Cancer Res.* 53:5740-5744). The *in vivo* mouse solid tumor xenograft model is used

-186-

in assays for that test an agent's ability to inhibit tumor cell proliferation and/or spontaneous metastasis. For example, a conjugate is evaluated for anti-tumor activity against any tumor subtype that expresses the targeted cell surface protease, *e.g.*, an ovarian tumor, in a mouse tumor  
5 xenograft model. Nude mice are given one or more, such as four intravenous injections of the conjugate. Dosing material is prepared by mixing the test material with appropriate volumes of, for example, PBS/0.1% BSA to achieve the desired doses. Mice IV injections (250-300 ul) into the tail vein for the duration of the experiment, such as, for  
10 example, days 5, 12, 19 and 26, with day 1 designated as the day that the tumor cells are injected into the mice. Doses are either fixed or normalized for differences in body weight. Tumor volume is measured twice weekly for a selected period.

Female Balb/c nu/nu athymic mice (Roger Williams Hospital Animal  
15 Facility, Providence, RI), 8-12 weeks old are suitable mice. They should be maintained in an aseptic environment and selected such that body weights range from about 25-30 grams the day prior to dosing. Animals are maintained in a quarantined room and handled under aseptic conditions. Food and water are supplied *ad libitum*. Appropriate tumor  
20 cells can be obtained, for example, from the American Type Culture Collection (Rockville, MD) and grown in modified Eagle's medium supplemented with 10% fetal calf serum. A selected number of days, such as five days prior to injection of the test material, mice receive a subcutaneous injection of tumor cells in the right rear flank.

25 Calipers are used to measure the dimensions of each tumor. Measurements (mm) of maximum and minimum width are performed prior to injection of the test material and at selected, such as bi-weekly, intervals for the duration of the experiment. Tumor volumes ( $\text{mm}^3$ ) can be computed, for example, using the formula:

-187-

Volume = [(width)<sup>2</sup>(length)]/2.

**G. Methods for Identifying proteases to target**

Also provided are methods for identifying proteases to target conjugates for treatment of diseases. The methods involve identifying  
5 cell-surface protease-associated disease by identifying a cell involved in the disease process or a cell in the vicinity of the cell involved in the disease process. For example, if disease involves a particular tumor, a protease present on the particular tumor or on cells that are located in the vicinity thereof is identified. A cell surface protease on the cell for  
10 targeting and substrates therefor are then identified. Conjugates that target such proteases as provided herein can then be prepared.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

15 **EXAMPLE 1**

**General Procedures for Preparing Peptide-doxorubicin conjugates**

**Step A: Synthesis of Peptides on Wang resin**

Peptides were prepared automatically using an ABI 431A peptide synthesizer from Perseptive Biosystems on preloaded Wang resin (0.25  
20 mmol). The ABI 431A uses HOBt, HBTU, DIEA activation. The synthesis of N-acetyl (or other amide) capped peptides involved the use of AcOH (or other respective carboxylic acid) during the final coupling step on the ABI 431A. Other N-terminal caps were attached manually by using the following reagents: For carbamates and sulfonamides the  
25 peptides were capped with ROCOCl or RSO<sub>2</sub>Cl and DIEA (4 equivalents each, 1 hr) in DMF (3 mL).

**Step B: Cleavage of peptides from Wang resin**

The cleavage of peptides from Wang resin involved shaking the resin with 2 mL TFA/H<sub>2</sub>O (95:5) for 45 min. The resin was removed by



-188-

- filtration and the filtrate was allowed to stand for an additional hour. The solution was concentrated to a residue. The crude peptide was analyzed by analytical HPLC (system A). Typical purity of the crude peptide ranged from 80% to 95%. The peptides were purified by preparative
- 5 HPLC (system B) using an appropriate gradient (typically 10-30%). Pure fractions were then lyophilized to provide the desired peptide as a white solid. Typical yields were 20-50% and a purity of 96-99%.

**Analytical HPLC conditions (System A)**

- Column: Chromolith RP-18e 4.6 mm x 100 mm from EM science
- 10 Gradient: 5-50% B in A over 6 min
- Flow Rate: 4 mL/min
- Solvent A: 0.1% TFA in water
- Solvent B: 0.1% TFA in acetonitrile
- Wavelength: 210 nm, 280 nm

**15 Preparative HPLC conditions (System B)**

- Column: Ultro 120 5 C18Q 150 x 20 mm from Peeke Scientific
- Gradient: 0-20%, or 10-30% or 20-40% B in A over 40 min
- Flow Rate: 18 mL/min
- Solvent A: 0.1% TFA in water
- 20 Solvent B: acetonitrile
- Wavelength: 214 nm

**Step C: Coupling of peptide acids to doxorubicin**

- To a mixture of peptide acid (0.052 mmol, 1.2 equivalents), doxorubicin hydrochloride (0.043 mmol, 25 mg), and HATU (0.0604
- 25 mmol, 22.9 mg, 1.4 equivalents) was added DMF (1 mL) then 2,6-lutidine (0.17 mmol, 20  $\mu$ L, 4 equivalents). The mixture was mixed until a homogeneous solution was obtained. After 4 to 24 hours (monitor by HPLC system A) the reaction was diluted with water (9 mL) and directly purified by preparative HPLC (system D). Pure fractions were then

-189-

lyophilized to provide the desired peptide doxorubicin conjugate as a fluffy red solid. The quality of the final conjugate was verified by analytical HPLC (system C) and mass spectroscopy. Typical yields were 10-30% with a purity of 95-99%. (Note: when the peptide acid

5 contained a histidine residue DIEA was substituted as the base and the reaction time was shortened to 1 hour).

Deprotection of fluorenylmethylesters of peptide doxorubicin conjugates: In cases where free carboxylic acid is present in the conjugate a fluorenyl methyl ester was used to protect a carboxylic acid

10 during coupling of the C-terminus of the peptide acid to doxorubicin, the fluorenylmethyl group was subsequently removed with 10% morpholine in DMF for 1 hour.

#### Analytical HPLC conditions (System C)

Column: Chromolith RP-18e 4.6 mm x 100 mm from EM science

15 Gradient: 5-50% B in A over 6 min

Flow Rate: 4 mL/min

Solvent A: 0.1% TFA in water

Solvent B: 0.1% TFA in acetonitrile

Wavelength: 210 nm, 280 nm

#### 20 Examples of retention times (min)

Doxorubicin	4.05
Ac-Gly-Ser-Gly-Arg-Ser-nLeu-Dox	4.34
MeOCO-Thr-Gly-Arg-Ser-nLeu-Dox	4.39
PhSO <sub>2</sub> -Thr-Gly-Arg-Ser-nLeu-Dox	4.83
25 N,N-dimethylglycine-Thr-Gly-Arg-Ser-nLeu-Dox	4.27
Ac-Thr-Gly-Arg-Ser-nLeu-Dox	4.32

#### Preparative HPLC conditions (System D)

Column: Ultro 120 5 C18Q 150 x 20 mm from Peeke Scientific

Gradient: 10-30% B in A over 40 min

-190-

Flow Rate: 18 mL/min

Solvent A: 0.1% acetic acid in water

Solvent B: acetonitrile

Wavelength: 214 nm

5

## EXAMPLE 2

### Preparation of Ac-Gly-Ser-Gly-Arg-Ser-nLeu-Dox

#### Step A: Manual synthesis of Ac-Gly-Ser(tBu)-Gly-Arg(Pbf)-Ser(tBu)-nLeu-Wang resin

- In a 250 mL fritted peptide synthesis vessel equipped with nitrogen
- 10 agitation and vacuum assisted drainage, Fmoc-nL-Wang resin (nova-biochem, 3.3 grams, 0.9 mmol/g, 3 mmol) was pre-swelled for 30 min using DMF. The peptide was then elongated by repeating the 4 step procedure below a total of five times with the following Fmoc aminoacids: Fmoc-Ser(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-
- 15 Ser(tBu)-OH, Fmoc-Gly-OH.

#### Iterative Coupling Procedure

1. the resin was mixed with 20% piperidine in DMF (100 mL) for 5 min then drained (repeat 3 times).
2. the resin was agitated with DMF (100 mL) for 30 sec then
- 20 drained (repeat 3 times).
3. to a mixture of Fmoc-aminoacid (12 mmol), HOBT (12 mmol, 4 equivalents, 1.622 g), TBTU (11.7 mmol, 3.9 equivalents, 3.757 g), DMF (10 mL) and NMP (90 mL) was added DIEA (12 mmol, 4 equivalents, 2.10 mL). After stirring for 5 min to allow pre-
- 25 activation, the solution was added to the synthesis vessel. The reaction was checked for completion by ninhydrin test and then drained. (If the ninhydrin test was blue, a double coupling (repeat step 3) was performed.
4. the resin was agitated with DMF (100 mL) for 30 sec then
- 30 drained (repeat 3 times).

-191-

The elongated resin (Fmoc-Gly-Ser(tBu)-Gly-Arg(Pbf)-Ser(tBu)-nLeu-Wang resin) was treated to steps 1 and 2 above to remove the Fmoc group. A solution of acetic anhydride (15 mmol, 5 equivalents, 1.42 mL), DIEA (15 mmol, 5 equivalents, 2.62 mL), DMF (10 mL) and NMP (90 mL) was added to the reaction vessel. After 1 hour the resin was washed with DMF (100 mL, 3 times), CH<sub>2</sub>Cl<sub>2</sub> (100 mL, 3 times) and MeOH (100 mL, 3 times). The resin was dried under vacuum for 15 hours.

**Step B: Preparation of Ac-Gly-Ser-Gly-Arg-Ser-nLeu-OH**

To the above synthesis vessel containing Ac-Gly-Ser(tBu)-Gly-Arg(Pbf)-Ser(tBu)-nLeu-Wang resin (3 mmol) was added TFA/H<sub>2</sub>O (95:5, 50 mL). After gently agitation for 45 min the cleavage solution was collected and the filtrate was allowed to stand for an additional 90 min. The solution was concentrated to a residue. The crude peptide was analyzed by analytical HPLC (system A, RT = 1.73, purity = 90%). The residue was dissolved in water (50 mL) and hexanes (10 mL) and mixed. The hexanes layer was removed and the aqueous layer bubbled with nitrogen to evaporate any remaining hexanes. The crude peptide was purified by preparative HPLC (system E). Pure fractions were then lyophilized to provide Ac-Gly-Ser-Gly-Arg-Ser-nLeu-OH (1.04 g, 1.68 mmol, 56%) as a white solid. The purity was evaluated by analytical HPLC (system A, RT = 1.73 min, 97% purity) and the constitution by mass spectroscopy (ion observed at 617.9).

**Preparative HPLC conditions (System E)**

Column: Waters Delta-Pak radial compression column, 15  $\mu$ m, 100A  
25 Gradient: 5-15% B in A over 40 min  
Flow Rate: 80 mL/min  
Solvent A: 0.1% acetic acid in water  
Solvent B: acetonitrile  
Wavelength: 214 nm

-192-

**Step C: Preparation of Ac-Gly-Ser-Gly-Arg-Ser-nLeu-Dox**

To a mixture of Ac-Gly-Ser-Gly-Arg-Ser-nLeu-OH (1.68 mmol, 1.04 g, 1.1 equivalents), doxorubicin hydrochloride (1.53 mmol, 887.8 mg), and HATU (1.76 mmol, 669.6 mg, 1.15 equivalents) was added DMF (40 mL) then 2,6-lutidine (6.12 mmol, 709  $\mu$ L, 4 equivalents). The solution was stirred for 18 hours. The reaction was diluted with water (100 mL), acidified with acetic acid (400  $\mu$ L) and purified immediately in three batches by preparative HPLC (system E). Each red colored fraction was analyzed by analytical HPLC (system F). Fractions of greater than 95% purity were then combined. The acetonitrile was removed under vacuum and the remaining solution was lyophilized to provide Ac-Gly-Ser-Gly-Arg-Ser-nLeu-Dox (0.682 mmol, 780 mg, 45%) as a fluffy red solid. The purity was evaluated by analytical HPLC (system F, RT = 3.51 min, 95% purity) and the constitution by mass spectroscopy (ion observed at 1143.5).

**Analytical HPLC conditions (System F)**

Column: Chromolith RP-18e 4.6 mm x 100 mm from EM science

Gradient: 20-40% B in A over 6 min

Flow Rate: 4 mL/min

20 Solvent A: 0.1% TFA in water

Solvent B: 0.1% TFA in acetonitrile

Wavelength: 210 nm, 280 nm

**Preparative HPLC conditions (System E)**

Column: Waters Delta-Pak radial compression column, 15  $\mu$ m, 100A

25 Gradient: 15-25% B in A over 40 min

Flow Rate: 80 mL/min

Solvent A: 0.1% acetic acid in water

Solvent B: acetonitrile

Wavelength: 214 nm

-193-

**EXAMPLE 3****General Procedures for Preparing Peptide-Taxol conjugates****Step A: Synthesis of Peptides on Wang resin**

See Example 1, Step A.

**5 Step B: Cleavage of peptides from Wang resin**

See Example 1, Step B.

**Step C: Coupling of peptide acids to 7-Gly-Taxol or 7-Ala-Taxol**

To a mixture of peptide acid (0.0121 mmol, 1.1 equivalents), 7-Gly-Taxol or 7-Ala-Taxol (0.011 mmol), and HATU (0.0154 mmol, 5.9 mg, 1.4 equivalents) was added DMF (0.3 mL) then 2,6-lutidine (0.044 mmol, 5.1  $\mu$ L, 4 equivalents). The mixture was mixed until a homogeneous solution was obtained. After 4 to 24 hours (monitor by HPLC system H) the reaction was diluted with water (9 mL) and directly purified by preparative HPLC (system I). Pure fractions were then

15 lyophilized to provide the desired peptide taxol conjugate as a fluffy white solid. The quality of the final conjugate was verified by analytical HPLC (system H) and mass spectroscopy. Typical yields were 30-50% with a purity of 96-99%.

**Analytical HPLC conditions (System H)**

20 Column: Chromolith RP-18e 4.6 mm x 100 mm from EM science  
Gradient: 5-90% B in A over 6 min  
Flow Rate: 4 mL/min  
Solvent A: 0.1% TFA in water  
Solvent B: 0.1% TFA in acetonitrile

25 Wavelength: 210 nm, 280 nm

**Examples of retention times (min)**

Ac-Gln-Ser-Arg-Ala-Ala-Taxol	2.86
Ac-Gln-Ser-Arg-Ser-Ala-Ala-Taxol	2.79
Ac-Ser-Gly-Arg-Ala-Ser-Ala-Taxol	2.87

-194-

Ac-Arg-Ser-Arg-Ala-Ala-Taxol 2.80

Ac-Ser-Gly-Arg-Ser-Ser-Ala-Taxol 2.81

**Preparative HPLC conditions (System I)**

Column: Ultro 120 5 C18Q 150 x 20 mm from Peeke Scientific

5 Gradient: 20-45% B in A over 40 min

Flow Rate: 18 mL/min

Solvent A: 0.1% TFA in water

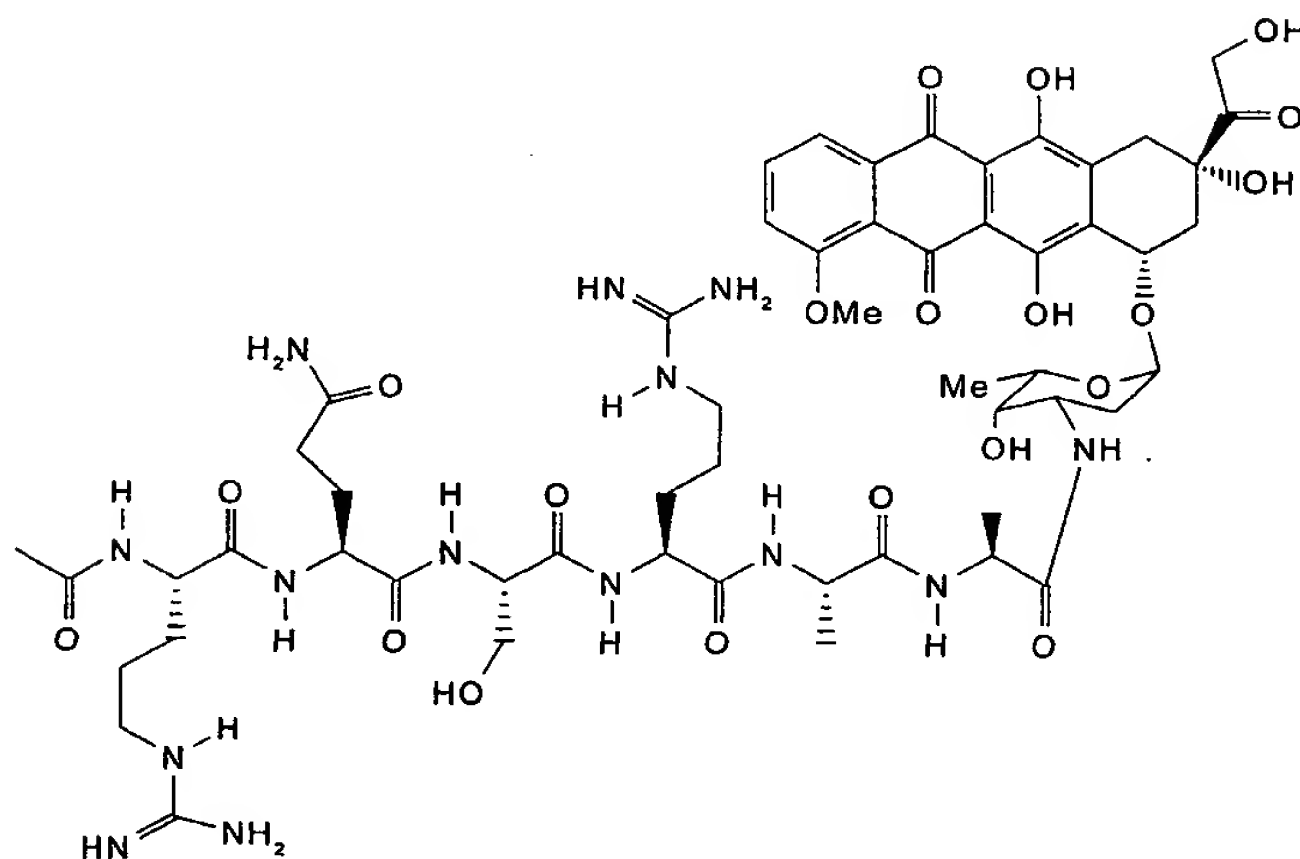
Solvent B: acetonitrile

Wavelength: 214 nm

10

**EXAMPLE 4**

**Preparation of N-Ac-Arg-Gln-Ser-Arg-Ala-Ala-DOX**



**Step A: N-Ac-Arg-Gln-Ser-Arg-Ala-Ala-OH**

Using the following general procedure, the *N*-acetyl peptidic  
 15 substrate *N*-Ac-Arg-Gln-Ser-Arg-Ala-Ala-OH was synthesized in a peptide  
 synthesis flask. Commencing with commercial Fmoc-Ala-Wang resin

-195-

(0.35 g, 0.84 mmol, Nova), standard Fmoc-deprotection with 20% piperidine was followed by a sequential iterative coupling-Fmoc deprotection strategy. Each coupling employed a 3-fold excess (2.52 mmol) of Fmoc-Ala, Fmoc-Arg(Boc)<sub>2</sub>, Fmoc-Ser(tBu), Fmoc-Gln(Trt) and  
5 Fmoc-Arg(Boc)<sub>2</sub>, respectively. Couplings were achieved using PyBOP (2.52 mmol) and DIEA (2.52 mmol) in DMF solvent. During each coupling cycle, the Fmoc protecting group was removed using 20% piperidine in DMF. After removal of the *N*-terminal Fmoc group, capping with acetic anhydride (1.43 mmol, 1.7 equiv.), DMAP (0.25 mmol, 0.3 equiv.), and  
10 DIEA (1.26 mmole, 1.5 equiv.) afforded the resin-bound *N*-acetyl intermediate. The protected peptide resin was treated with 50% TFA in methylene chloride for 30 min to cleave the Wang resin and then the Boc, Trt and t-Bu protecting groups were removed with 70% TFA in methylene chloride. Solvent and other volatile byproducts were evaporated under  
15 reduced pressure and the crude product was dissolved in water and lyophilized to afford the title compound as a nearly colorless, amorphous solid. Mass spectral analysis confirmed the desired molecular weight. HPLC analysis indicated the product to be of approximately 95% purity. The peptide carboxylic acid intermediate can be further purified by  
20 trituration or by preparative HPLC, if desired.

**Step B: N-Ac-Arg-Gln-Ser-Arg-Ala-Ala-DOX**

The intermediate from Step A (20 mg, 0.027 mmol) was dissolved in dry DMF (0.8 mL) and was stirred at room temperature under a nitrogen atmosphere. To this solution was added doxorubicin  
25 hydrochloride (15.6 mg, 0.027 mmol), EDC (6.8 mg, 0.035 mmol), HOAt (4.8 mg, 0.035 mmol) and 2,6-lutidine (7.3  $\mu$ L, 0.06 mmol). Stirring was continued until completion of the coupling, as monitored by analytical HPLC (system J, see below). The solution was filtered and the crude product was purified by C18 RP-HPLC (A=0.1% AcOH/H<sub>2</sub>O; B=CH<sub>3</sub>CN),



-196-

gradient elution 100% to 60% A over 60 min). Homogeneous product fractions (evaluated by HPLC, system J) were pooled and lyophilized to afford the title compound as a light red solid.

**HPLC conditions, system J:**

- 5 Column: Phenomenex 15 cm #00F-3033-E0, C18  
 Eluant: Gradient 95:5 (A:B) to 25:75 (A:B) over 20 min.  
 A = 0.1% TFA/H<sub>2</sub>O, B = 0.1% TFA/Acetonitrile  
 Flow: 1 mL/min.  
 Wavelength: 210 nm, 280 nm  
 10 Retention times: Doxorubicin = 8.89 min.  
 N-Ac-Arg-Gln-Ser-Arg-Ala-Ala-Dox = 8.4 min.

**Physical Properties:**

Molecular Formula: C<sub>55</sub>H<sub>78</sub>N<sub>14</sub>O<sub>20</sub>

Molecular Weight: 1255.3

- 15 Low Resolution Mass Spec: 628.2 (M + 2/2)

Table 2 lists data for additional peptidic substrate-Doxorubicin conjugates. These conjugates were prepared from the appropriate amino acid precursors that were elaborated by the general procedures described in Example 4.

20

**TABLE 2**

Peptidic substrate-DOX Conjugate	Mass Spectrum	HPLC-Retention Time (min.)
Acetyl-Arg-Arg-Gln-Ser-Arg-Ala-Ala-DOX	471.2 (M + 3/3)	8.23
25 Acetyl-Leu-Arg-Arg-Gln-Ser-Arg-Ala-Ala-DOX	509.2 (M + 3/3)	8.60

-197-

**EXAMPLE 5****Determination of times to 50% cleavage of Doxorubicin/peptidic substrate Conjugates by the recombinant protease domain of MTSP1**

One millimolar stock solutions were prepared for each peptidic  
5 substrate conjugate in double distilled water. Cleavage reactions were  
then performed in which 100  $\mu$ M conjugate was mixed with 1 or 10 nM  
of the recombinantly-produced active single chain protease domain of  
MTSP1 (residue 615-855 in SEQ ID No. 2, encoded by nucleotides 1865-  
2582 in SEQ ID No. 1) in 29.2 mM Tris, pH 8.4, 29.2 mM Imidazole, 217  
10 mM NaCl. Final reaction volume was 200  $\mu$ L. These reactions were  
incubated in a water bath at 37 °C. At times ranging from 2 to 128  
minutes, 20  $\mu$ L samples were removed, and enzymatic activity was  
stopped by the addition of trifluoroacetic acid to 0.33%. The amount of  
hydrolysis in each sample was measured by reverse phase HPLC. Percent  
15 hydrolysis was then calculated by dividing the area under the product  
peak by the sum of the areas under substrate and product peaks.  
Percent unhydrolyzed substrate was plotted against log of reaction times,  
and the plots were fit to sigmoidal curves using Prism software from  
Graphpad Inc. (San Diego, CA) to determine times at which 50% of each  
20 substrate was cleaved.

Results for certain of the conjugates provided herein are shown in  
Figure 1 (conditions: 1 nM MTSP1 with 100  $\mu$ M conjugate at 37 °C in  
12 mM tris(hydroxymethyl)aminomethane, pH 8.0, 25 mM NaCl, 0.5 mM  
CaCl<sub>2</sub>; reactions were quenched with 0.33% trifluoroacetic acid).

**25 EXAMPLE 6*****In vitro* assay of cytotoxicity of Conjugates**

The cytotoxicity of the conjugates also can be tested to confirm  
that the conjugates act as prodrugs. The conjugates are tested against a  
line of cells, which is known to be killed by unmodified cytotoxic agent,  
30 using an Alamar Blue assay. Cells, such as LNCaP cells (The American

-198-

Type Culture Collection (Rockville, Maryland)), that express a cell surface protease, such as MTSP1 or endotheliase, are seeded in 96 well plates at a density of  $1 \times 10^4$  cells/well (0.1 mL/well). A plate containing medium alone is used as a control. The cells are incubated for 3 days at 37 °C and 20  $\mu$ L of Alamar Blue is added to the assay well(s). After 7 h of incubation, cell killing is measured using an EL-310 plate reader at 570 and 600 nm. Values for cell killing are expressed as the percentage reduction in cell numbers relative to the media controls.

#### EXAMPLE 7

##### 10 *In vivo* efficacy of Conjugates

Tumor cells are trypsinized, resuspended in the growth medium and centrifuged for 6 min at 200xg. The cells are resuspended in serum-free  $\alpha$ -MEM and counted. The appropriate volume of this solution containing the desired number of cells is then transferred to a conical centrifuge tube, centrifuged as before and resuspended in the appropriate volume of a cold 1:1 mixture of  $\alpha$ -MEM-Matrigel. The suspension is kept on ice until the animals are inoculated.

Male nude mice 10 weeks of age are used. Mice are individually weighed and assigned to groups ( $n = 10$  per group) with no more than a 2-gram difference in weight between individual mice within each group. On day 1, mice are inoculated subcutaneously with the tumor cell line. Each mouse is inoculated with, for example, 0.5 mL of  $0.5 \times 10^6$  to  $10^8$  tumor cells/mL in a 60% solution of ice-cold Matrigel and  $\alpha$ -MEM. Then, 24 h later, conjugate administration began. Vehicle-treated mice are injected with 5% dextrose in water. At the end of a predetermined time, such as 18 days to two months or more, the mice are sacrificed, and tumor size and mass or other parameters are measured. Tumor size and mass or the other parameters for conjugate-treated mice are compared to vehicle-treated mice to determine efficacy of the conjugate.

-199-

Following inoculation with the tumor cells the mice are treated under one of three protocols:

**Protocol A**

One day after cell inoculation the animals are dosed with 1 to 100,  
5 or 3 to 50, or 5 to 25, or 7 to 22  $\mu\text{mol/kg}$ , including 7.2 or 17.9  $\mu\text{mol/kg}$ , of test conjugate, unmodified cytotoxic agent or vehicle control (sterile water). Dosages of the conjugate and cytotoxic agent are initially the maximum non-lethal amount, but can be subsequently titrated lower. Identical doses are administered at 24 hour intervals for 5 days. At the  
10 end of 5.5 weeks or other suitable interval, the mice are sacrificed and weights of any tumors present are measured. The animals' weights are determined at the beginning and end of the assay.

**Protocol B**

At 14-15 days after cell inoculation, the animals are dosed with 1  
15 to 100, or 3 to 50, or 5 to 25, or 7 to 22  $\mu\text{mol/kg}$ , including 7.2 or 17.9  $\mu\text{mol/kg}$ , of test conjugate, unmodified cytotoxic agent, or vehicle control (sterile water). Dosages of the conjugate and cytotoxic agent are initially the maximum non-lethal amount, but can be subsequently titrated lower. Identical doses are administered at 24 hour intervals for 5 days. At the  
20 end of 5.5 weeks or other suitable interval, the mice are sacrificed and weights of any tumors present are measured. The animals' weights are determined at the beginning and end of the assay.

**Protocol C**

One day after cell inoculation, the animals are dosed by  
25 interperitoneal administration with 1 to 100, or 3 to 50, or 5 to 25, or 7 to 22  $\mu\text{mol/kg}$ , including 7.2 or 17.9  $\mu\text{mol/kg}$ , of test conjugate, unmodified cytotoxic agent, or vehicle control (sterile water). Dosages of the conjugate and cytotoxic agent are initially the maximum non-lethal amount, but can be subsequently titrated lower. Identical doses are

-200-

administered at 7 day intervals for 5 weeks. At the end of 5.5 weeks or other suitable interval, the mice are sacrificed and weights of any tumors present are measured. The animals' weights are determined at the beginning and end of the assay.

5 **EXAMPLE 8**

**Gene expression profiles of exemplary MTSPs and Domain organization**

**Gene expression profile of MTSP1 in normal tissues, cancer cells and cancer tissues**

To obtain information regarding the tissue distribution and gene  
10 expression level of MTSP1, the DNA insert from a *Pichia pastoris*  
expression vector, pPIC9K-MTSP1, containing the encoding nucleic acid,  
was used to probe a blot containing RNA from 76 different human tissues  
(catalog number 7775-1; human multiple tissue expression (MTE) array;  
CLONTECH, Palo Alto, CA). Significant expression was observed in the  
15 colon (ascending, transverse and descending), rectum, trachea,  
esophagus and duodenum. Moderate expression levels were observed in  
the jejunum, ileum, ileocecum, stomach, prostate, pituitary gland,  
appendix, kidney, lung, placenta, pancreas, thyroid gland, salivary gland,  
mammary gland, fetal kidney, and fetal lung. Lower expression levels  
20 were seen in the spleen, thymus, peripheral blood leukocyte, lymph node,  
bone marrow, bladder, uterus, liver, adrenal gland, fetal heart, fetal liver,  
fetal spleen, and fetal thymus. A significant amount of the MTSP1  
transcript was also detected in colorectal adenocarcinoma cell line  
(SW480), Burkitt's lymphoma cell line (Daudi), and leukemia cell line (HL-  
25 60). RT-PCR of the MTSP1 transcript in several human primary tumors  
xenografted in athymic nude mice was performed using gene-specific  
primers. A high level of MTSP1 transcript was detected in colon  
adenocarcinoma (CX-1) and pancreatic adenocarcinoma (GI-103).  
Moderate levels were observed in another colon adenocarcinoma  
30 (GI-112), ovarian carcinoma (GI-102), lung carcinoma (LX-1), and breast

-201-

carcinoma (GI-101). Another lung carcinoma (GI-117) expressed a low level of the MTSP1 transcript. A similar RT-PCR was performed to detect the presence of the MTSP1 transcript in PC-3 and LNCaP cell lines. Both cell lines expressed significant amounts of MTSP1 transcript. MTSP1  
5 also is a marker for ovarian cancer.

**Gene expression profile of the serine protease MTSP3 in normal and tumor tissues**

To obtain information regarding the tissue distribution of the  
10 MTSP3 transcripts, the DNA insert encoding the MTSP3 protease domain was used to probe a RNA blot composed of 76 different human tissues (catalog number 7775-1; human multiple tissue expression (MTE) array; CLONTECH, Palo Alto, CA). The expression pattern observed in decreasing signal level was: trachea = colon (descending) = esophagus  
15 > colon (ascending) > colon (transverse) = rectum > ileum > duodenum > jejunum > bladder > ileocecum > stomach > kidney > appendix. It also is expressed less abundantly in fetal kidney, and in two tumor cell lines, HeLa S3 and leukemia, K-562. Northern analysis using RNA blots (catalog numbers 7780-1, 7765-1 & 7782-1; human 12-lane,  
20 human muscle and human digestive system multiple tissue northern (MTN) blots; CLONTECH) confirmed that the expression was detected most abundantly in the colon, moderately in the esophagus, small intestine, bladder and kidney, and less abundantly in stomach and rectum. A single transcript of ~2.2 kb was detected.

25 , Amplification of the MTSP3 transcript in several human primary tumors xenografted in mouse was performed using gene-specific primers. The MTSP3 transcript was detected in lung carcinoma (LX-1), colon adenocarcinoma (CX-1), colon adenocarcinoma (GI-112) and ovarian carcinoma (GI-102). No apparent signal was detected in another form of

-202-

lung carcinoma (GI-117), breast carcinoma (GI-101), pancreatic adenocarcinoma (GI-103) and prostatic adenocarcinoma (PC3).

**Gene expression profile of MTSP4 in normal and tumor tissues**

- To obtain information regarding the gene expression profile of the
- 5 MTSP4 transcript, a DNA fragment encoding part of the LDL receptor domain and the protease domain was used to probe an RNA blot composed of 76 different human tissues (catalog number 7775-1; human multiple tissue expression (MTE) array; CLONTECH). As in the northern analysis of gel blot, a very strong signal was observed in the liver.
- 10 Signals in other tissues were observed in (decreasing signal level): fetal liver > heart = kidney = adrenal gland = testis = fetal heart and kidney = skeletal muscle = bladder = placenta > brain = spinal cord = colon = stomach = spleen = lymph node = bone marrow = trachea = uterus = pancreas = salivary gland = mammary gland = lung. MTSP4 also is
- 15 expressed less abundantly in several tumor cell lines including HeLa S3 = leukemia K-562 = Burkitt's lymphomas (Raji and Daudi) = colorectal adenocarcinoma (SW480) > lung carcinoma (A549) = leukemia MOLT-4 = leukemia HL-60. PCR of the MTSP4 transcript from cDNA libraries made from several human primary tumors xenografted in nude mice
- 20 (human tumor multiple tissue cDNA panel, catalog number K1522-1, CLONTECH) was performed using MTSP4-specific primers. The MTSP4 transcript was detected in breast carcinoma (GI-101), lung carcinoma (LX-1), colon adenocarcinoma (GI-112) and pancreatic adenocarcinoma (GI-103). No apparent signal was detected in another form of lung
- 25 carcinoma (GI-117), colon adenocarcinoma (CX-1), ovarian carcinoma (GI-102). and prostatic adenocarcinoma (PC3). The MTSP4 transcript was also detected in LNCaP and PC-3 prostate cancer cell lines as well as in HT-1080 human fibrosarcoma cell line.

-203-

**Gene expression profile of MTSP6 in normal and tumor tissues**

To obtain information regarding the gene expression profile of the MTSP6 transcript, a 495 bp DNA fragment obtained from PCR reaction with primers Ch17-NSP-3 and NSP-4AS was used to probe an RNA blot

5 composed of 76 different human tissues (catalog number 7775-1; human multiple tissue expression (MTE) array; CLONTECH). The strongest signal was observed in duodenum. Signals in other tissues were observed in (decreased signal level): Stomach > trachea = mammary gland = thyroid gland = salivary gland = pituitary gland = pancreas > kidney >

10 lung > jejunum = ileum = ileocecum = appendix = fetal kidney > fetal lung. Very weak signals also can be detected in several other tissues. MTSP6 also is expressed in several tumor cell lines including HeLa S3 > colorectal adenocarcinoma (SW480) > leukemia MOLT-4 > leukemia K-562. PCR analysis of the MTSP6 transcript from cDNA libraries made

15 from several human primary tumors xenografted in nude mice (human tumor multiple tissue cDNA panel, catalog number K1522-1, CLONTECH) was performed using MTSP6-specific primers (Ch17-NSP-3 and Ch17-NSP2AS). The MTSP6 transcript was strongly detected in lung carcinoma (LX-1), moderately detected in pancreatic adenocarcinoma

20 (GI-103), weakly detected in ovarian carcinoma (GI-102); and very weakly detected in colon adenocarcinoma (GI-112 and CX-1), breast carcinoma (GI-101), lung carcinoma (GI-117) and prostatic adenocarcinoma (PC3). The MTSP6 transcript was also detected in breast cancer cell line MDA-MB-231, prostate cancer cell line PC-3, but

25 not in HT-1080 human fibrosarcoma cell line. MTSP6 also is expressed in mammary gland carcinoma cDNA (Clontech).

**Gene expression profile of MTSP9 in normal, tumor tissues and cell lines**

To obtain a gene expression profile of the MTSP9 transcript, the

30 MTSP9 cDNA fragment obtained from human pancreas was used to



-204-

probe a dot blot composed of RNA extracted from 76 different human tissues (Human Multiple Tissue Expression (MTE) Array; Clontech, Palo Alto, CA; catalog no. 7775-1). The results of this analysis indicate that MTSP9 is highly expressed in the esophagus and expressed at a low level

5 in many other tissues. The MTSP9 transcript is found in kidney (adult and fetal), spleen (adult and fetal), placenta, liver (adult and fetal), thymus, peripheral blood leukocyte, lung (adult and fetal), pancreas, lymph node, bone marrow, trachea, uterus, prostate, esophagus, testes, ovary and the gland organs (mammary, adrenal, thyroid, pituitary and

10 salivary). MTSP9 also is expressed in tumor esophagus tissues, in a lung carcinoma (A549 cell line) and, at a low level, in a colorectal carcinoma (SW480), lymphoma (Raji and Daudi), a cervical carcinoma (HeLaS3) and leukemia (HL-60, K-562 and MOLT-4) cell lines.

**Gene expression profile of MTSP10 in normal and tumor tissues**

15 To obtain information regarding the gene expression profile of the MTSP10 transcript, PCR analysis was carried out on cDNA panels made from several human adult tissues (Clontech, Cat. #K1420-1) cDNA panel using MTSP10-specific primers. MTSP10 transcript was detected in pancreas, lung and kidney. MTSP10 transcript was also detected in small

20 intestine Marathon-Ready cDNA (Clontech). PCR of the MTSP10 transcript from cDNA libraries made from several human primary tumors xenografted in nude mice (human tumor multiple tissue cDNA panel, catalog number K1522-1, CLONTECH) was also performed. The MTSP10 transcript was detected in breast carcinoma (GI-101), lung

25 carcinoma (LX-1 and GI-117), ovarian carcinoma (GI-102), and pancreatic adenocarcinoma (GI-103). The MTSP10 transcript can be weakly detected in prostatic adenocarcinoma (PC3). No apparent signal was detected in two forms of colon adenocarcinomas (GI-112 and CX-1). The

-205-

MTSP10 transcript was also detected in CWR22R prostate tumor grown on nude mice.

**Domain organization and gene expression profile of MTSP12 in normal and tumor tissues**

**5                    Domain organization of MTSP12PD1, -PD2 and -PD3 and homology to other serine proteases**

Sequence and protein domain analyses of the translated MTSP12PD1, -PD2 and -PD3 nucleotide sequences indicate that these three serine proteases are contiguous. The sequence order is as follows:

10 MTSP12-PD1 is found at the N terminus followed by MTSP12-PD2, and MTSP12-PD3 is at the C terminus. MTSP12-PD1 and -PD2 contain a trypsin-like serine protease domain (aa 236 to aa 465 and aa 537 to aa 765 for MTSP12-PD1 and -PD2, respectively) characterized by the presence of a protease activation cleavage site (...R<sub>236</sub> ↓ I<sub>237</sub>VGGMEAS...,

15 and ... R<sub>537</sub> ↓ V<sub>538</sub>VGGFGAA..., for MTSP12-PD1 and -PD2, respectively, and where | indicates a protease activation cleavage site) and the catalytic triad residues (His<sub>277</sub>, Asp<sub>326</sub> and Ser<sub>421</sub> in MTSP12-PD1; His<sub>578</sub>, Asp<sub>626</sub> and Ser<sub>721</sub> in MTSP12-PD2) in 3 highly-conserved regions of the catalytic domain. MTSP12-PD3 contains a serine protease domain (aa

20 861 to aa 1087); it has a protease activation cleavage site (...R<sub>860</sub> ↓ I<sub>861</sub>VGGSAAG...) and has the catalytic His<sub>902</sub> and Asp<sub>949</sub>, but it has a Ala<sub>1043</sub> instead of the conserved catalytic serine found in serine proteases. Several domains are found upstream of the MTSP12-PD1 serine protease domain and these include a transmembrane domain (aa

25 28 to aa 50), a SEA (sea urchin sperm protein-enterokinase-agrin) domain (aa 51 to aa 170) and an LDLa (low density lipoprotein receptor class a) domain (aa 187 to aa 225). There are 5 possible *N*-linked glycosylation sites (N<sub>116</sub>SS, N<sub>581</sub>HT, N<sub>672</sub>AT, N<sub>697</sub>FS and N<sub>820</sub>ST). In the protease domain of MTSP12-PD1, there is an unpaired cysteine (C<sub>346</sub>) in a single

30 chain form of the protease domain and the following Cys pairings are

-206-

noted: C<sub>262</sub>-C<sub>278</sub>; C<sub>360</sub>-C<sub>427</sub>; C<sub>417</sub>-C<sub>446</sub>; C<sub>392</sub>-C<sub>406</sub>. In the protease domain of MTSP12-PD2, there is an unpaired cysteine (C<sub>646</sub>) in a single chain form of the protease domain, and the following Cys pairings are noted: C<sub>563</sub>-C<sub>579</sub>; C<sub>660</sub>-C<sub>727</sub>; C<sub>692</sub>-C<sub>706</sub>; C<sub>717</sub>-C<sub>746</sub>. In the protease domain of MTSP12-PD3, there is an unpaired cysteine (C<sub>969</sub>) in a single chain form of the protease domain, and the following Cys pairings are noted: C<sub>887</sub>-C<sub>903</sub>; C<sub>983</sub>-C<sub>1049</sub>; C<sub>1014</sub>-C<sub>1028</sub>; C<sub>1039</sub>-C<sub>1068</sub>.

Alignment (*blastp*; <http://www.ncbi.nlm.nih.gov/BLAST>) of the respective MTSP12-PD1, MTSP12-PD2 and MTSP12-PD3 protein sequences to known serine proteases deposited in the public database showed a 45%, 45% and 48% identity to matriptase, a 44%, 43% and 41% identity with DESC1/endotheliasin 1, a 44%, 43% and 48% identity to prostamin (AB030036), a 43%, 39% and 39% identity to spinesin (TMPRSS5; NM\_030770), and a 40%, 38% and 38% identity to marapsin (NM\_031948). The clone has about 93% homology at the nucleotide and encoded protein levels to a clone and encoded provided described in International PCT application No. WO 02/00860 (see SEQ ID Nos. 38 and 97 therein). The encoded protein described in the PCT application, however, includes the Sequence set forth in SEQ ID No. 271 between amino acids Leu373 and Val374 of SEQ ID No. 20, as well as an additional extended sequence of amino acids between amino acids Ala48 and Phe49 of SEQ ID No. 20 and lacks amino acids 91-124 of SEQ ID No. 20. The protein provided in International PCT application No. WO02/00860 can be used in the methods provided herein.

#### 25                                      **Gene and Tissue expression profile of MTSP12**

To obtain information regarding the tissue distribution profile of the MTSP12PD1, -PD2 and -PD3 transcripts, 3 cDNA probes were prepared. Data indicate that the MTSP12PD1, -PD2 and -PD3 transcript is

-207-

expressed at a low level in most of the 76 tissues and cell lines, but at a higher level in the lymph node and testes.

To compare the expression profile of MTSP12PD1, -PD2 and -PD3 in a range of normal human and matched tumor tissues, a matched  
5 tumor/normal expression array (catalog number 7840-1; <http://www.clontech.com>) composed of 68 paired cDNA samples from individual patients was used. Results show that the MTSP12PD1, -PD2 and -PD3 transcript is expressed at a low level in a number of normal tissues including breast, uterus, colon, ovary, lung, kidney and rectum,  
10 but is not differentially expressed in any of the matched tumors. It also is expressed at a low level in several tumor cell lines, including HeLa (cervical carcinoma), Daudi (Burkitt's lymphoma), K562 (chronic myelogenous leukemia), HL-60 (premyelocytic leukemia), G361 (melanoma), A549 (lung carcinoma), MOLT-4 (lymphoblastic leukemia),  
15 SW480 (colorectal adenocarcinoma), and Raji (Burkitt's lymphoma).

Several SMART™ 5'-RACE cDNA libraries (catalog number K1811-1; <http://www.clontech.com>) prepared from normal breast, normal testes, normal prostate, prostate cancer cell lines and breast cancer cell lines were analyzed for the presence of MTSP12PD1, -PD2 and -PD3 transcript  
20 by RT-PCR using two sets of gene-specific primers. The MTSP12-PD2 and -PD3 transcript was detected in normal prostate, PC-3, LNCaP, normal breast, MDA-MB-231, MDA-MB-361, MDA-MB-453 and DU4475, but higher levels were observed in normal breast and MDA-MB-231. The MTSP12-PD1 transcript was detected in the same tissues and cell lines,  
25 except higher levels were observed in normal breast, MDA-MB-231 and DU4475.

-208-

**Gene expression profile of MTSP20 in normal, tumor tissues and cell lines**

To obtain information regarding the gene expression profile of the MTSP20 transcript, the MTSP20 cDNA fragment obtained from human lung tissue was used to probe a dot blot composed of RNA extracted from 76 different human tissues (Human Multiple Tissue Expression (MTE) Array; Clontech, Palo Alto, CA; catalog no. 7775-1). The results indicate that RNA encoding MTSP20 is expressed in a variety of tissues. The MTSP20 transcript is found in liver, lymph node, cerebellum, pancreas, prostate, uterus, testis, glands (adrenal, thyroid and salivary), thymus, kidney and spleen. Lower transcript level can be found in lung, placenta, bladder, ovary, digestive system, circulatory system and other parts of the the brain. MTSP20 is also expressed in certain tumor cell lines including lung carcinoma (A519), colorectal carcinoma (SW480), lymphoma (Raji and Daudi), cervical carcinoma (HeLaS3) and leukemia (HL-60, K-562 and MOLT-4) cell lines.

**Gene expression profile of MTSP22 in normal, tumor tissues and cell lines**

MTSP22 is expressed in the uterine tissue, thymus, adipose tissue, and lymph node. It may also be expressed in lung, stomach, uterine, breast, ovarian, prostate and in other tumors. To obtain information regarding the gene expression profile of the MTSP22 transcript, the cDNA fragment encoding the entire serine protease domain was used to probe a dot blot composed of RNA extracted from 72 different human tissues (Human Multiple Tissue Expression (MTE) Array; Clontech, Palo Alto, CA; catalog no. 7776-1) as well as a dot blot composed of normalized cDNA from 241 tumor and corresponding normal tissues from individual patients (Cancer Profiling Array, Clontech, catalog no. 7841-1). The results of MTE analysis indicated that MTSP22 transcript is expressed primarily in the esophagus. In the cancer profiling array analysis, MTSP22 is

-209-

highly expressed in 3 of the 42 normal uterus tissue samples, but not in their matched tumor samples. In one of the 42 uterus samples, MTSP22 is expressed in tumor and its metastatic tissues, but not in the normal tissue counterpart. MTSP22 is also expressed in 2 of the 17 stomach  
5 tumors and 2 of the 21 lung tumors, but not in their normal tissue counterparts. MTSP22 is also expressed in the normal tissue of the only pancreas matched cDNA pair. PCR analysis was also performed using commercially available cDNA panel from several human adult tissues (Clontech, Cat. #K1420-1 and K1420-2) and primary tumors (Clontech  
10 Cat. #K1522-1) as well as several Marathon-Ready cDNAs (Clontech).

MTSP22 cDNA was detected in thymus, adipose tissue, and lymph node. Serine protease domain of MTSP22 and homology to other proteases.

Sequence analysis of the translated MTSP22 protease domain  
15 sequence revealed that MTSP22 contains a trypsin-like serine protease domain characterized by the presence of a protease activation cleavage site at the amino terminus of the domain and the catalytic triad residues (histidine, aspartate and serine) in three highly-conserved regions. Alignment of the protein sequence with that of endotheliase 1 (same as  
20 serine protease DESC1 protein; GenBank accession number AF064819) indicated that the two proteins share 50% sequence identity in their protease domains.

**Gene expression profile of MTSP25 in normal, tumor tissues and cell lines**

25 MTSP25 is expressed in breast, colon, uterine, ovarian, kidney, prostate, testicular cancer tissue. It may also be expressed in lung, stomach, prostate and in other tumors. To obtain information regarding the gene expression profile of the MTSP25 transcript, a 369 bp DNA fragment containing MTSP25 protease domain sequence (obtained from a  
30 PCR reaction) was used to probe a dot blot composed of RNA extracted

-210-

from 72 different human tissues (Human Multiple Tissue Expression (MTE) Array; Clontech, Palo Alto, CA; catalog no. 7776-1) as well as a dot blot composed of normalized cDNA from 241 tumor and corresponding normal tissues from individual patients (Cancer Profiling Array, Clontech, catalog no. 7841-1). The results of MTE analysis indicate that MTSP25 transcript is expressed weakly in the lymph node. In the cancer profiling array analysis, MTSP25 is highly expressed in all 4 prostate samples (in normal and cancer samples). In one of the 20 kidney cDNA pairs, MTSP25 is highly expressed in the tumor sample, but not in its normal tissue counterpart. MTSP25 is also expressed in 1 of the 50 breast cancer samples, but not in its normal tissue counterpart.

MTSP25 is also expressed in 3 of the 42 normal uterus samples, but not in their tumor counterparts. MTSP25 expression is also detected in 3 of the 14 ovarian cancer samples. Among these three samples, the expression of MTSP25 was also detected in one of the matched normal tissue counterparts. MTSP25 expression was also detected in 5 tumor samples in the 34 colon cDNA pairs.

PCR analysis was also performed using a commercially available cDNA panel from several human adult tissues (Clontech, Cat. #K1420-1 and K1420-2) as well as several Marathon-Ready cDNAs (Clontech). MTSP25 cDNA was strongly detected in testis and mammary gland adenocarcinoma, weakly detected in brain, placenta, lung, spleen, prostate, small intestine, colon, and leukocyte, and very weakly detected in heart, liver, and pancreas.

25

#### EXAMPLE 9

Conjugates that have been prepared according to the procedures of Examples 1-4 by routine and minor modification of the procedures, such as using different Fmoc-amino acid building blocks, include:

Ac-R-Q-G-R-S-L-(Dox) (SEQ ID NO: 491);

-211-

- Ac-R-Q-G-R-S-S-L-(Dox) (SEQ ID NO: 492);  
Ac-R-Q-G-R-S-nL-(Dox) (SEQ ID NO: 493);  
Ac-R-Q-G-R-S-nV-(Dox) (SEQ ID NO: 494);  
Ac-R-Q-G-R-S-F-(Dox) (SEQ ID NO: 495);  
5 Ac-R-Q-G-R-A-L-(Dox) (SEQ ID NO: 496);  
Ac-R-Q-G-R-A-L-(Dox) (SEQ ID NO: 497);  
Ac-R-Q-G-R-A-nL-(Dox) (SEQ ID NO: 498);  
Ac-R-Q-G-R-A-nL-(Dox) (SEQ ID NO: 499);  
Ac-R-Q-G-R-A-nV-(Dox) (SEQ ID NO: 500);  
10 Ac-R-Q-G-R-A-Cha-(Dox) (SEQ ID NO: 501);  
Ac-R-Q-G-R-A-F-(Dox) (SEQ ID NO: 502);  
Ac-R-N-G-R-S-L-(Dox) (SEQ ID NO: 503);  
Ac-R-N-G-R-A-nL-(Dox) (SEQ ID NO: 504);  
Ac-R-Q-A-R-S-L-(Dox) (SEQ ID NO: 505);  
15 Ac-R-Q-A-R-S-nL-(Dox) (SEQ ID NO: 506);  
Ac-R-Q-A-R-S-nV-(Dox) (SEQ ID NO: 507);  
Ac-R-Q-A-A-S-Cha-(Dox) (SEQ ID NO: 508);  
Ac-R-Q-A-R-S-S-Cha-(Dox) (SEQ ID NO: 509);  
Ac-R-Q-A-R-T-nL-(Dox) (SEQ ID NO: 510);  
20 Ac-R-Q-A-R-A-L-(Dox) (SEQ ID NO: 511);  
Ac-R-Q-A-R-A-nL-(Dox) (SEQ ID NO: 513);  
Ac-R-Q-A-R-A-nV-(Dox) (SEQ ID NO: 514);  
Ac-R-Q-A-R-A-Cha-(Dox) (SEQ ID NO: 515);  
Ac-R-Q-S-R-A-A-(Dox) (SEQ ID NO: 516);  
25 Ac-R-Q-S-R-A-(Dox) (SEQ ID NO: 517);  
Ac-R-Q-S-R-A-nL-(Dox) (SEQ ID NO: 518);  
Ac-R-Q-S-R-A-L-(Dox) (SEQ ID NO: 519);  
Ac-R-Q-S-R-A-nV-(Dox) (SEQ ID NO: 520);  
Ac-R-Q-S-R-A-Cha-(Dox) (SEQ ID NO: 521);



-212-

- Ac-R-Q-S-R-S-S-L-(Dox) (SEQ ID NO: 522);  
Ac-R-Q-S-R-S-L-(Dox) (SEQ ID NO: 523);  
Ac-R-Q-S-R-S-dnL-(Dox) (SEQ ID NO: 524);  
Ac-R-Q-S-R-S-nL-(Dox) (SEQ ID NO: 525);  
5 Ac-R-Q-S-R-S-nV-(Dox) (SEQ ID NO: 526);  
Ac-R-Q-S-R-S-allylG-(Dox) (SEQ ID NO: 527);  
Ac-R-Q-S-R-S-Cha-(Dox) (SEQ ID NO: 528);  
Ac-R-Q-S-R-T-nL-(Dox) (SEQ ID NO: 529);  
Ac-R-Q-T-R-S-S-L-(Dox) (SEQ ID NO: 530);  
10 Ac-R-Q-T-R-S-L-(Dox) (SEQ ID NO: 531);  
Ac-R-N-S-R-S-nL-(Dox) (SEQ ID NO: 532);  
Ac-R-Q-F-R-S-L-(Dox) (SEQ ID NO: 533);  
Ac-R-Q-F-R-S-nL-(Dox) (SEQ ID NO: 534);  
Ac-R-Q-F-R-S-nV-(Dox) (SEQ ID NO: 535);  
15 Ac-R-Q-F-R-S-nL-(Dox) (SEQ ID NO: 536);  
Ac-R-Q-F-R-S-Cha-(Dox) (SEQ ID NO: 537);  
Ac-R-Q-F-R-A-L-(Dox) (SEQ ID NO: 538);  
Ac-R-Q-F-R-A-nL-(Dox) (SEQ ID NO: 539);  
Ac-R-Q-F-R-A-nV-(Dox) (SEQ ID NO: 540);  
20 Ac-R-Q-F-R-A-Cha-(Dox) (SEQ ID NO: 541);  
Ac-Q-S-R-S-S-nL-(Dox) (SEQ ID NO: 542);  
MeOCO-Quat2-G-R-S-L-NH2 (SEQ ID NO: 483);  
MeOCO-Quat3-G-R-S-L-NH2 (SEQ ID NO: 484);  
MeOCO-Quat-G-R-S-L-NH2 (SEQ ID NO: 485);  
25 MeOCO-Quat4-G-R-S-L-NH2 (SEQ ID NO: 486);  
MeOCO-Quat5-G-R-S-L-NH2 (SEQ ID NO: 487);  
MeOCO-Quat2-G-R-S-S-L-NH2 (SEQ ID NO: 488);  
MeOCO-Quat4-G-R-S-L-(Dox) (SEQ ID NO: 489);  
MeOCO-Quat2-G-R-S-L-(Dox) (SEQ ID NO: 490);

-213-

- Ac-Q-G-R-S-L-(Dox) (SEQ ID NO: 445);  
Ac-Q-G-R-S-S-L-(Dox) (SEQ ID NO: 446);  
Ac-Q-G-R-A-S-L-(Dox) (SEQ ID NO: 447);  
Ac-N-G-R-S-S-L-(Dox) (SEQ ID NO: 448);  
5 Ac-Q-G-R-S-S-nL-(Dox) (SEQ ID NO: 449);  
Ac-Q-G-R-S-S-nV-(Dox) (SEQ ID NO: 450);  
Ac-Q-G-R-S-S-Cha-(Dox) (SEQ ID NO: 451);  
Ac-Q-G-R-S-S-allylG-(Dox) (SEQ ID NO: 452);  
Ac-Q-G-R-S-S-allylG-(Dox) (SEQ ID NO: 453);  
10 Ac-Q-A-R-S-L-(Dox) (SEQ ID NO: 454);  
Ac-Q-A-R-S-S-L-(Dox) (SEQ ID NO: 455);  
Ac-Q-S-R-S-L-(Dox) (SEQ ID NO: 456);  
Ac-Q-S-R-S-S-nV-(Dox) (SEQ ID NO: 457);  
Ac-Q-S-R-S-S-Cha-(Dox) (SEQ ID NO: 458);  
15 Ac-Q-S-R-S-S-L-(Dox) (SEQ ID NO: 459);  
Ac-Q-T-R-S-S-L-(Dox) (SEQ ID NO: 460);  
Ac-Q-Aib-R-S-S-Cha-(Dox) (SEQ ID NO: 461);  
Ac-Q-Aib -R-S-S-L-(Dox) (SEQ ID NO: 462);  
Ac-Q-Abu-R-S-S-Cha-(Dox) (SEQ ID NO: 463);  
20 Ac-Q-Abu-R-S-S-L-(Dox) (SEQ ID NO: 464);  
Ac-Q-Cha-R-S-S-Cha-(Dox) (SEQ ID NO: 465);  
Ac-Q-F-R-S-L-(Dox) (SEQ ID NO: 466);  
Ac-Q-F-R-S-S-L-(Dox) (SEQ ID NO: 467);  
Ac-Q-Y-R-S-S-L-(Dox) (SEQ ID NO: 468);  
25 Ac-R-G-R-S-L-(Dox) (SEQ ID NO: 469);  
Ac-R-G-R-S-S-L-(Dox) (SEQ ID NO: 470);  
Ac-R-G-R-S-S-Cha-(Dox) (SEQ ID NO: 471);  
Ac-R-G-R-S-Cha-(Dox) (SEQ ID NO: 472);  
Ac-R-A-R-S-L-(Dox) (SEQ ID NO: 473);

-214-

- Ac-R-A-R-S-S-L-(Dox) (SEQ ID NO: 474);  
Ac-R-S-R-S-L-(Dox) (SEQ ID NO: 475);  
Ac-R-S-R-S-S-L-(Dox) (SEQ ID NO: 476);  
Ac-R-S-R-S-Cha-(Dox) (SEQ ID NO: 477);  
5 Ac-R-S-R-S-S-Cha-(Dox) (SEQ ID NO: 478);  
Ac-R-F-R-S-L-(Dox) (SEQ ID NO: 479);  
Ac-R-F-R-S-Cha-(Dox) (SEQ ID NO: 480);  
Ac-Y-G-R-S-S-L-(Dox) (SEQ ID NO: 481);  
Ac-M(O2)-S-R-S-L-(Dox) (SEQ ID NO: 482);  
10 Ac-R-R-Q-S-R-A-A-(Dox) (SEQ ID NO: 105);  
Ac-R-R-Q-S-R-I-(Dox) (SEQ ID NO: 610);  
Ac-R-R-Q-S-R-S-S-L-(Dox) (SEQ ID NO: 543);  
Ac-R-R-Q-S-R-S-L-(Dox) (SEQ ID NO: 544);  
Ac-R-G-S-G-R-S-L-(Dox) (SEQ ID NO: 545);  
15 Ac-R-G-S-G-R--S-nL-(Dox) (SEQ ID NO: 546);  
Ac-R-G-S-G-R-A-nL-(Dox) (SEQ ID NO: 547);  
Ac-R-G-S-G-R-S-S-L-(Dox) (SEQ ID NO: 548);  
Ac-I-V-S-G-R-A-S-L-(Dox) (SEQ ID NO: 549);  
Ac-R-R-Q-S-R-A-(Dox) (SEQ ID NO: 108);  
20 Ac-R-R-Q-S-R-I-(Dox) (SEQ ID NO: 111);  
Ac-L-R-R-Q-S-R-A-A-(Dox) (SEQ ID NO: 106);  
Ac-L-R-R-Q-S-R-G-G-(Dox) (SEQ ID NO: 109);  
Ac-L-R-R-Q-S-R-A-(Dox) (SEQ ID NO: 110);  
Ac-L-R-R-Q-S-R-A-I-(Dox) (SEQ ID NO: 112);  
25 Ac-L-R-R-Q-S-R-A-I-(Dox) (SEQ ID NO: 611);  
Ac-L-R-R-Q-S-R-S-S-L-(Dox) (SEQ ID NO: 550);  
Ac-L-R-R-Q-S-R-S-L-(Dox) (SEQ ID NO: 551);  
Ac-S-G-R-S-L-(Dox) (SEQ ID NO: 362);  
Ac-S-G-R-S-S-L-(Dox) (SEQ ID NO: 363);

-215-

- Ac-S-G-R-S-S-S-L-(Dox) (SEQ ID NO: 364);  
Ac-S-G-R-S-nL-(Dox) (SEQ ID NO: 365);  
Ac-S-G-R-S-nV-(Dox) (SEQ ID NO: 366); isomer 1  
Ac-S-G-R-S-nV-(Dox) (SEQ ID NO: 367); isomer 2  
5 Ac-S-G-R-S-G(hex)-(Dox) (SEQ ID NO: 368);  
Ac-S-G-R-S-Cha-(Dox) (SEQ ID NO: 369);  
Ac-S-G-R-S-hCha-(Dox) (SEQ ID NO: 370);  
Ac-S-A-R-S-L-(Dox) (SEQ ID NO: 371);  
Ac-S-A-R-S-S-L-(Dox) (SEQ ID NO: 372);  
10 Ac-S-S-R-S-nL-(Dox) (SEQ ID NO: 373);  
Ac-T-G-R-S-Abu-(Dox) (SEQ ID NO: 374);  
Ac-T-G-R-S-L-(Dox) (SEQ ID NO: 375);  
Ac-T-G-R-S-nV-(Dox) (SEQ ID NO: 376);  
Ac-T-G-R-S-nL-(Dox) (SEQ ID NO: 377);  
15 Ac-T-G-R-S-G(hex)-(Dox) (SEQ ID NO: 378);  
Ac-T-G-R-S-Cha-(Dox) (SEQ ID NO: 379);  
Ac-T-G-R-S-hCha-(Dox) (SEQ ID NO: 380);  
Ac-T-G-R-T-Abu-(Dox) (SEQ ID NO: 381);  
Ac-T-G-R-hS-nL-(Dox) (SEQ ID NO: 382);  
20 Ac-T-G-R-Abu-nL-(Dox) (SEQ ID NO: 383);  
Ac-T-G-R-Abu-nV-(Dox) (SEQ ID NO: 384);  
Ac-T-G-F(Gn)-S-nL-(Dox) (SEQ ID NO: 385);  
Ac-T-G-F(Gn)-S-Cha-(Dox) (SEQ ID NO: 386);  
Ac-T-G-F(Gn)-Abu-nV-(Dox) (SEQ ID NO: 387);  
25 Ac-T-G-K(alloc)-S-nL-(Dox) (SEQ ID NO: 388);  
Ac-T-G-K-S-nL-(Dox) (SEQ ID NO: 389);  
Ac-T-G-hR-S-nL-(Dox) (SEQ ID NO: 390);  
Ac-(hS)G-G-R-S-nL-(Dox) (SEQ ID NO: 391);  
MeOCO-T-G-R-S-nL-(Dox) (SEQ ID NO: 392);

-216-

- PhSO<sub>2</sub>-T-G-R-S-nL-(Dox) (SEQ ID NO: 393);  
MeOEtCO-T-G-R-S-nL-(Dox) (SEQ ID NO: 394);  
MeO(EtO)<sub>2</sub>Ac-T-G-R-S-nL-(Dox) (SEQ ID NO: 395);  
4-oxo-Pentanoyl-T-G-R-S-nL-(Dox) (SEQ ID NO: 396);  
5 3,4-MethyldioxyPhAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 397);  
2-PyridylAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 398);  
PhOAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 399);  
L-3-PhLactyl-T-G-R-S-nL-(Dox) (SEQ ID NO: 400);  
MeOAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 401);  
10 PhAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 402);  
MeOEtOCO-T-G-R-S-nL-(Dox) (SEQ ID NO: 403);  
MeOEtOAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 404);  
HOOCButa-T-G-R-S-nL-(Dox) (SEQ ID NO: 405);  
Z-T-G-R-S-nL-(Dox) (SEQ ID NO: 406);  
15 EtOCO-T-G-R-S-nL-(Dox) (SEQ ID NO: 407);  
 $\beta$ A-T-G-R-S-nL-(Dox) (SEQ ID NO: 408);  
Pent-4-ynoyl-T-G-R-S-nL-(Dox) (SEQ ID NO: 409);  
NapAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 410);  
iBoc-T-G-R-S-nL-(Dox) (SEQ ID NO: 411);  
20 HOAc-T-G-R-S-nL-(Dox) (SEQ ID NO: 412);  
MeSucc-T-G-R-S-nL-(Dox) (SEQ ID NO: 413);  
N,N-diMeGly-T-G-R-S-nL-(Dox) (SEQ ID NO: 414);  
Succ-T-G-R-S-nL-(Dox) (SEQ ID NO: 415);  
HCO-T-G-R-S-nL-(Dox) (SEQ ID NO: 416);  
25 Ac-T-A-R-S-nL-(Dox) (SEQ ID NO: 417);  
Ac-T-A-F(Gn)-S-nL-(Dox) (SEQ ID NO: 418);  
Ac-T-A-R-Abu-nV-(Dox) (SEQ ID NO: 419);  
Ac-T-A-R-S-Abu-(Dox) (SEQ ID NO: 420);  
Ac-T-A-R-T-Abu-(Dox) (SEQ ID NO: 421);

-217-

- Ac-T-S(O-Me)-R-S-nL-(Dox) (SEQ ID NO: 422);  
Ac-T-hS-R-S-nL-(Dox) (SEQ ID NO: 423);  
Ac-T-(1-Me)H-R-S-nL-(Dox) (SEQ ID NO: 424);  
Ac-T-(3-Me)H-R-S-nL-(Dox) (SEQ ID NO: 425);  
5 Ac-T-H-R-S-nL-(Dox) (SEQ ID NO: 426);  
Ac-T-Sar-R-S-nL-(Dox) (SEQ ID NO: 427);  
Ac-T-nV-R-S-nL-(Dox) (SEQ ID NO: 428);  
Ac-T-nL-R-S-nL-(Dox) (SEQ ID NO: 429);  
Ac-T-A-R-S-Cha-(Dox) (SEQ ID NO: 430);  
10 Ac-T-Abu-R-S-nL-(Dox) (SEQ ID NO: 431);  
Ac-4,4diMeThr-G-R-S-nL-(Dox) (SEQ ID NO: 432);  
Ac-hS-G-R-S-nL-(Dox) (SEQ ID NO: 433);  
Ac-hS-G-R-hS-Cha-(Dox) (SEQ ID NO: 434);  
Ac-hS-G-R-S-Cha-(Dox) (SEQ ID NO: 435);  
15 Ac-hS-G-R-T-Cha-(Dox) (SEQ ID NO: 436);  
Ac-hS-A-R-S-Cha-(Dox) (SEQ ID NO: 437);  
Ac-N-G-R-S-nL-(Dox) (SEQ ID NO: 438);  
Ac-Y-G-R-S-S-L-(Dox) (SEQ ID NO: 439);  
Ac-Y-G-R-S-Cha-(Dox) (SEQ ID NO: 440);  
20 Ac-Q-G-R-S-S-nL-(Dox) (SEQ ID NO: 441);  
Ac-Q-G-R-S-S-nV-(Dox) (SEQ ID NO: 442);  
Ac-L-R-G-S-G-R-S-A-(Dox) (SEQ ID NO: 573);  
Ac-L-R-G-S-G-R-S-L-(Dox) (SEQ ID NO: 342);  
Ac-L-R-G-S-G-R-S-L-(Dox) (SEQ ID NO: 343);  
25 Ac-L-R-G-S-G-R-S-S-nL-(Dox) (SEQ ID NO: 344);  
Ac-L-R-G-S-G-R-S-S-Cha-(Dox) (SEQ ID NO: 345);  
Ac-L-R-G-dS-A-R-S-A-(Dox) (SEQ ID NO: 574);  
Ac-L-R-G-S-A-R-S-S-L-(Dox) (SEQ ID NO: 346);  
Ac-L-R-G-S-A-R-S-L-(Dox) (SEQ ID NO: 347);

-218-

- Ac-L-R-G-S-A-R-S-S-Cha-(Dox) (SEQ ID NO: 348);  
Ac-L-R-G-S-A-R-S-S-nV-(Dox) (SEQ ID NO: 349);  
Ac-L-R-G-S-A-R-S-S-nL-(Dox) (SEQ ID NO: 350);  
Ac-V-I-V-S-G-R-A-L-(Dox) (SEQ ID NO: 351);  
5 Ac-V-I-V-S-A-R-S-L-(Dox) (SEQ ID NO: 352);  
Ac-V-I-V-S-G-R-S-S-L-(Dox) (SEQ ID NO: 353);  
Ac-V-I-V-S-A-R-M-A-(Dox) (SEQ ID NO: 354);  
Ac-V-I-V-S-A-R-nL-A-(Dox) (SEQ ID NO: 355);  
Ac-V-I-V-S-A-R-S-nL-(Dox) (SEQ ID NO: 356);  
10 Ac-V-I-V-S-A-R-S-Cha-(Dox) (SEQ ID NO: 357);  
Ac-V-I-V-S-A-R-S-Cha-(Dox) (SEQ ID NO: 358);  
Ac-V-I-V-S-A-R-S-S-Cha-(Dox) (SEQ ID NO: 359);  
Ac-R-R-(Me)C-P-G-R-V-V-(Dox) (SEQ ID NO: 360);  
Ac-R-R-nV-P-A-R-S-L-(Dox) (SEQ ID NO: 361);  
15 Ac-R-G-dS-A-R-S-A-(Dox) (SEQ ID NO: 309);  
Ac-R-G-S-G-R-S-A-(Dox) (SEQ ID NO: 310);  
Ac-R-G-S-G-R-A-L-(Dox) (SEQ ID NO: 311);  
Ac-R-G-S-G-R-S-L-(Dox) (SEQ ID NO: 312);  
Ac-R-G-S-G-R--S-nL-(Dox) (SEQ ID NO: 313);  
20 Ac-R-G-S-G-R-A-nL-(Dox) (SEQ ID NO: 314);  
Ac-R-G-S-G-R-S-S-L-(Dox) (SEQ ID NO: 315);  
Ac-R-G-S-G-R-S-Cha-(Dox) (SEQ ID NO: 316);  
Ac-R-G-S-G-R-S-S-Cha-(Dox) (SEQ ID NO: 317);  
Ac-R-G-S-A-R-S-Cha-(Dox) (SEQ ID NO: 318);  
25 Ac-R-G-S-A-R-S-S-(Dox) (SEQ ID NO: 319);  
Ac-R-G-S-A-R-S-nV-(Dox) (SEQ ID NO: 320);  
Ac-R-G-S-A-R-S-S-nV -(Dox) (SEQ ID NO: 321);  
Ac-R-G-S-A-R-S-L-(Dox) (SEQ ID NO: 322);  
Ac-R-(Me)C-P-G-R-V-V-(Dox) (SEQ ID NO: 323);

-219-

- Ac-R-(Me)C-P-G-R-V-V-(Dox) (SEQ ID NO: 324);  
Ac-R-C(Me)-P-G-R-S-L-(Dox) (SEQ ID NO: 325);  
Ac-R-L-P-G-R-S-L-(Dox) (SEQ ID NO: 326);  
Ac-R-V-P-G-R-S-L-(Dox) (SEQ ID NO: 327);  
5 Ac-R-V-P-G-R-S-L-(Dox) (SEQ ID NO: 328);  
Ac-R-nL-P-G-R-S-L-(Dox) (SEQ ID NO: 329);  
Ac-R-G(tBu)-P-A-R-S-L-(Dox) (SEQ ID NO: 330);  
Ac-R-L-P-A-R-S-L-(Dox) (SEQ ID NO: 331);  
Ac-R-V-P-A-R-S-L-(Dox) (SEQ ID NO: 332);  
10 Ac-R-nL-P-A-R-S-L-(Dox) (SEQ ID NO: 333);  
Ac-I-V-S-G-R-A-L-(Dox) (SEQ ID NO: 334);  
Ac-I-V-S-G-R-S-S-L-(Dox) (SEQ ID NO: 335);  
Ac-I-V-S-G-R-A-S-L-(Dox) (SEQ ID NO: 336);  
Ac-I-V-S-A-R-M-A-(Dox) (SEQ ID NO: 337);  
15 Ac-I-V-S-A-R-nL-A-(Dox) (SEQ ID NO: 338);  
Ac-I-V-S-A-R-S-L-(Dox) (SEQ ID NO: 339);  
Ac-I-V-S-A-R-S-nL-(Dox) (SEQ ID NO: 340);  
Ac-I-V-S-A-R-S-S-L-(Dox) (SEQ ID NO: 341);  
Ac-G-S-G-R-S-A-(Dox) (SEQ ID NO: 585);  
20 Ac-G-S-G-R-S-L-(Dox) (SEQ ID NO: 277);  
Ac-G-S-G-R-A-L-(Dox) (SEQ ID NO: 278);  
Ac-G-S-G-R-S-S-L-(Dox) (SEQ ID NO: 279);  
Ac-G-S-G-R-L-(Dox) (SEQ ID NO: 280);  
Ac-G-S-G-(4-guan)Phg-S-L-NH<sub>2</sub> (SEQ ID NO: 281);  
25 Ac-G-S-G-R-S-S-Cha-(Dox) (SEQ ID NO: 282);  
Ac-G-S-G-R-A-S-L-(Dox) (SEQ ID NO: 283);  
Ac-G-S-G-R-S-nL-(Dox) (SEQ ID NO: 284);  
Ac-G-T-G-R-S-nL-(Dox) (SEQ ID NO: 285);  
Succ-bA-T-G-R-S-nL-(Dox) (SEQ ID NO: 286);



-220-

- Ac-G-T-G-R-S-hCha-(Dox) (SEQ ID NO: 287);  
Ac-G-hS-G-R-S-nL-(Dox) (SEQ ID NO: 288);  
Ac-G-dS-A-R-S-A-(Dox) (SEQ ID NO: 289);  
Ac-G-S-A-R-S-L-(Dox) (SEQ ID NO: 290);  
5 Ac-G-S-A-R-S-S-Cha-(Dox) (SEQ ID NO: 291);  
Ac-G-S-A-R-S-S-L-(Dox) (SEQ ID NO: 292);  
Ac-G-S-A-R-A-S-L-(Dox) (SEQ ID NO: 293);  
Ac-V-S-G-R-S-L-(Dox) (SEQ ID NO: 294);  
Ac-V-S-G-R-A-L-(Dox) (SEQ ID NO: 295);  
10 Ac-V-S-G-R-A-S-L-(Dox) (SEQ ID NO: 296);  
Ac-V-S-G-R-S-S-L-(Dox) (SEQ ID NO: 297);  
Ac-V-S-A-R-M-A-(Dox) (SEQ ID NO: 298);  
Ac-V-S-A-R-nL-A-(Dox) (SEQ ID NO: 299);  
Ac-V-S-A-R-S-nL-(Dox) (SEQ ID NO: 300);  
15 Ac-V-S-A-R-S-L-(Dox) (SEQ ID NO: 301);  
Ac-(Me)C-P-G-R-V-V-(Dox) (SEQ ID NO: 302);  
Ac-(Me)C-P-G-R-V-V-(Dox) (SEQ ID NO: 303);  
Ac-C(Me)-P-G-R-A-L-(Dox) (SEQ ID NO: 304);  
Ac-C(Me)-P-G-R-S-L-(Dox) (SEQ ID NO: 305);  
20 Ac-C(Me)-P-A-R-S-L-(Dox) (SEQ ID NO: 306);  
Ac-C(Me)-P-A-R-A-S-L-(Dox) (SEQ ID NO: 307);  
Ac-G(tBu)-P-G-R-S-L-(Dox) (SEQ ID NO: 308);  
Ac-Q-S-R-A-A-(taxol) (SEQ ID NO: 552);  
Ac-Q-S-R-S-A-(taxol) (SEQ ID NO: 553);  
25 Ac-Q-S-R-S-G-(taxol) (SEQ ID NO: 554);  
Ac-R-S-R-A-A-(taxol) (SEQ ID NO: 555);  
Ac-R-Q-S-R-A-A-(taxol) (SEQ ID NO: 556);  
Ac-R-Q-S-R-S-A-(taxol) (SEQ ID NO: 557);  
Ac-R-Q-S-R-S-A-A-(taxol) (SEQ ID NO: 558);

-221-

- Ac-R-G-S-G-R-S-A-(taxol) (SEQ ID NO: 559);  
 Ac-S-G-R-A-A-(taxol) (SEQ ID NO: 560);  
 Ac-S-G-R-S-A-(taxol) (SEQ ID NO: 561);  
 Ac-S-G-R-S-S-A-(taxol) (SEQ ID NO: 562);  
 5 Ac-S-G-R-A-S-A-(taxol) (SEQ ID NO: 563);  
 Ac-S-G-R-S-G-(taxol) (SEQ ID NO: 564);  
 Ac-S-G-R-S-S-G-(taxol) (SEQ ID NO: 565);  
 Ac-S-G-R-S-G-A-(taxol) (SEQ ID NO: 566);  
 Ac-S-G-R-S-G-G-(taxol) (SEQ ID NO: 567);  
 10 Ac-G-T-G-R-S-G-G-(taxol) (SEQ ID NO: 568);  
 Ac-L-R-R-Q-S-R-A-A-(Dox) (SEQ ID NO: 597);  
 MeSO<sub>2</sub>-dA(Chx)-Abu-R-S-L-(Dox) (SEQ ID NO: 598);  
 Ac-R-A-R-S-L-(Dox) (SEQ ID NO: 599);  
 Ac-dA(Chx)-Abu-R-S-L-(Dox) (SEQ ID NO: 600);  
 15 Ac-dA(Chx)-Abu-R-S-S-L-(Dox) (SEQ ID NO: 601);  
 Ac-Q-G-R-S-S-L-(Dox) (SEQ ID NO: 602);  
 MeOCO-dhF-P(OH)-R-S-S-L-(Dox) (SEQ ID NO: 603);  
 MeOCO-Quat4-G-R-S-L-(Dox) (SEQ ID NO: 604);  
 As-dCha-P(OH)-R-S-S-L-(Dox) (SEQ ID NO: 605);  
 20 Ac-dCha-Abu-R-S-S-A-(taxol) (SEQ ID NO: 606);  
 MeOCO-Quat2-G-R-S-L-NH<sub>2</sub> (SEQ ID NO: 607);  
 MeOCO-Quat3-G-R-S-L-NH<sub>2</sub> (SEQ ID NO: 608); and  
 MeOCO-Quat-G-R-S-L-NH<sub>2</sub> (SEQ ID NO: 609).

#### EXAMPLE 10

- 25 **Pharmacokinetic studies of conjugates and fraction of the dose metabolized to Doxorubicin and Leucine-doxorubicin in naïve and tumor bearing mice.**

Naïve or tumor bearing nude mice 8-12 weeks of age have been used for pharmacokinetic studies of the test conjugates. Tumor cells for  
 30 implantation have been prepared following one of three protocols.

-222-

**Protocol A Tumor cells collected from tissue culture**

Tumor cells are trypsinized and resuspended in the growth medium and centrifuged for 6 min at 200xg. The cells are resuspended in serum-free medium and counted. The appropriate volume of the solution

- 5 containing the desired number of cells is then transferred to a conical centrifuge tube, centrifuged as before and resuspended in the appropriate volume of a cold 1:1 mixture of cells in phenol free medium : matrigel. Each mouse is inoculated with 0.2 - 0.5 mL containing between  $1 \times 10^6$  and  $1 \times 10^7$  tumor cells subcutaneously or orthotopically.

**10 Protocol B Tumor cell suspension**

Established tumors (200-1000mm<sup>3</sup>) are dissected from mice, weighed and rinsed in tumor cell growth medium. The tumors are passed through a steel cell dissociation sieve. The cells are rinsed through the sieve with growth medium. The cells are centrifuged for 6 min at 200xg

15 and resuspended in the appropriate volume of a cold 1:1 mixture of cells : matrigel. Each mouse is inoculated with 0.2-0.5 mL of tumor cells subcutaneously or orthotopically.

**Protocol C Tumor fragments**

- Alternatively a tumor measuring approximately 800mm<sup>3</sup> is
- 20 dissected out of a mouse, rinsed in tumor cell growth medium and cut into 1-2 mm<sup>3</sup> fragments. Each fragment is inoculated subcutaneously or orthotopically using a trocar needle.

**Pharmacokinetic Study**

- Naïve or tumor bearing mice are individually weighed and assigned
- 25 to groups. The mice are dosed with 1-100umole/kg, including 30umole/kg, 25umole/kg, or 21.5umole/kg of the test conjugate intraperitoneally or intravenously. At a given time point between 5 minutes and 24 hours after administration of the compound the mice are sacrificed. Blood is collected in a syringe containing protease inhibitors

-223-

such as EDTA, AEBSF, Aprotinin, Leupeptin, Bestatin, Pepstatin A or E64 and transferred into a heparinized blood collection tube. The plasma is prepared by centrifugation. The tumors are collected and pulverized in liquid nitrogen. The resulting tumor powders are stored at  $-80^{\circ}\text{C}$ . The  
5 tumor powders and plasma are extracted and analyzed for the parent test conjugate and its products including Leucine-doxorubicin (or norleucine-doxorubicin, etc.) and doxorubicin.

Looking at the delivery of the toxin to the tumor cells, and also looking at the parent conjugate and the levels of toxin (dox and nor-leu  
10 dox) in the plasma.

### RESULTS

For example, test conjugate (21.5 umole/kg of Ac-Gly-Ser-Gly-Arg-Ser-nLeu-Dox (see Example 2)) was administered to naïve and tumor bearing (TB) mice intraperitoneally (IP) or intravenously (IV). One hour  
15 after administration plasma and tumor tissue was collected from the mice. Concentrations of the test conjugate and its products are compared. The results show that the conjugate does not get into the tumor, the toxins (norleu dox and dox -  $\mu\text{M}$  concentrations in the tumor at one hour following the single (both IP and IV) injection. There were lower  
20 levels of dox and nor-leu dox the plasma than in the tumor.

#### Extraction, chromatography LC/MS conditions

**Plasma:** Plasma samples are prepared using acetonitrile protein precipitation. A standard curve was constructed from addition of 5 to 20  $\mu\text{L}$  volumes of a standard compound to 0.1 mL or 0.05 mL volumes of  
25 plasma on ice. The standard curve ranges from 10 ng/mL – 1 ug/mL or from 100 ng/mL – 4 ug/mL of the standard compound. Immediately after standard addition, acetonitrile is added to precipitate the proteins. The study plasma samples were prepared by thawing the frozen plasma samples on ice. The aliquots were added directly to the acetonitrile.

-224-

After sample precipitation, the sample is mixed using vortex mixing. The precipitate was pelleted using centrifugation. The supernatant was dried using vacuum centrifugation. The sample was reconstituted with 0.15 mL of 30% acetonitrile – 70% (0.01 M ammonium acetate with 0.1% formic acid). 0.01 mL of the sample was injected for LC-MS analysis. The HPLC conditions were a linear gradient of 20% acetonitrile – 80% (10 mM ammonium acetate - .1% formic acid) to 50% acetonitrile – 50% (10 mM ammonium acetate - .1% formic acid) in 1 minute at 0.3 mL/min in a 30 x 2.1 mm Zorbax SB C18 HPLC column. Detection was provided by a triple quad mass spectrometer with electrospray ionization. Doxorubicin was monitored using the m/z transition 544.1 – 396.8. Leucine-doxorubicin was monitored using 657.2 – 242.8. An exemplary parent conjugate was monitored using 1555.9 – 1555.9. Scanning LC-MS and fluorescence detection was used to identify cleavage products other than doxorubicin or leucine-doxorubicin (or norleucine-doxorubicin, etc.) in the plasma.

**Tumor:** Immediately after excision from the mouse, the tumor for analysis is weighed and placed into a mortar containing liquid nitrogen. With the mortar nested in a bed of dry ice, the tumor is ground into a fine powder while additional liquid nitrogen is added as needed to avoid thawing. When a homogeneous tumor powder is achieved, the remaining liquid nitrogen is allowed to boil off. The tumor powder is quantitatively transferred to a 15ml conical tube that has been pre-chilled and is on dry ice. The sample is stored at –70 °C until analysis. The tumor powder is thawed on ice and vortex mixed with 0.01M ammonium acetate in a 1 gram tumor/mL ammonium acetate solution concentration to form a slurry. An aliquot of 0.1mL of the tumor slurry is precipitated with 0.5 mL acetonitrile. The supernatant is separated from the precipitated solids and then evaporated using vacuum centrifugation. Quantification of

-225-

doxorubicin, leucine-doxorubicin (or norleucine-doxorubicin, etc.), is achieved by reference to a standard curve constructed from spiking measured amounts of standard compounds (doxorubicin, leucine-doxorubicin, etc.) into control tumor slurry. A typical standard curve  
5 ranges from 1 ng to 200 ng of compound per aliquot of tumor slurry. After the unknown samples and standards are processed and dried, the residue is reconstituted in 0.15mL of 30% acetonitrile – 70% (0.01M ammonium acetate + 0.1% formic acid). 10  $\mu$ L of solution is injected onto a liquid chromatography – mass spectrometry system. The HPLC  
10 conditions were a linear gradient of 20% acetonitrile – 80% (10 mM ammonium acetate - .1% formic acid) to 50% acetonitrile – 50% (10 mM ammonium acetate - .1% formic acid) in 1 minute at 0.3 mL/min in a 30 x 2.1 mm Zorbax SB C18 HPLC column. Detection was provided by a triple quad mass spectrometer with electrospray ionization. Doxorubicin  
15 was monitored using the m/z transition 544.1 – 396.8. Leucine-doxorubicin was monitored using 657.2 – 242.8.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended  
20 claims.

-226-

**WHAT IS CLAIMED IS:**

1. A conjugate, comprising a therapeutic agent and a peptidic substrate linked thereto optionally via a linker, wherein the peptidic substrate is proteolytically cleaved by a cell surface protease or a soluble,  
5 released or shed form thereof, to liberate the therapeutic agent, wherein the conjugate is not substantially cleaved by plasmin or prostate specific antigen (PSA).
2. The conjugate of claim 1, wherein the liberated therapeutic agent is active.
- 10 3. The conjugate of claim 1, wherein cleavage liberates the therapeutic agent in a form that requires further processing for activation.
4. The conjugate of claim 1 that comprises the components: (peptidic substrate)<sub>s</sub>, (Linker)<sub>q</sub>, and (therapeutic agent)<sub>t</sub>;  
wherein at least one peptidic substrate moiety is linked with or  
15 without a linker to at least one therapeutic agent, s is 1 to 6, q is 0 to t, and t is 1 to 6, wherein a cell surface protease that cleaves the peptidic substrate(s) results in delivery of the therapeutic agent to the cell.
5. The conjugate of claim 1, wherein the peptidic substrate comprises one amino acid or more, wherein, upon proteolytic cleavage of  
20 the conjugate, the resulting therapeutic agent is active or in a form that, upon further processing, is active.
6. The conjugate of claim 1, wherein the cell surface protease is a serine protease.
7. The conjugate of claim 1, wherein the cell surface protease  
25 is a type II transmembrane serine protease (MTSP) or an endotheliase.
8. The conjugate of claim 1, wherein the cell surface protease is selected from endotheliase 1, endotheliase 2, MTSP1, MTSP3, MTSP4, MTSP6, MTSP7, MTSP9, MTSP10, MTSP12, MTSP20, MTSP22, MTSP25, corin, enterokinase, human airway trypsin-like protease (HAT),

-227-

TMPRSS2, hepsin, urokinase-type plasminogen activator (uPA), and TMPRSS4.

9. The conjugate of claim 1, wherein the cell surface protease comprises a polypeptide selected from the group consisting of

5 a polypeptide comprising the sequence of amino acids set forth in any of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 29, 31, 33, 35, 37, 39, 41, 43, 45, 270, 272, 274 and 276;

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of high stringency to the sequence of  
10 nucleotides set forth in any of SEQ ID Nos 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 28, 30, 32, 34, 36, 38, 40, 42, 44, 269, 273 and 275;

a polypeptide that comprises a sequence of amino acids having at least about 40% sequence identity with the sequence of amino acids set forth in SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22,  
15 24, 26, 29, 31, 33, 35, 37, 39, 41, 43, 45, 270, 272, 274 and 276;  
and

a polypeptide encoded by a splice variant of the sequence of nucleotides set forth in any of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 29, 31, 33, 35, 37, 39, 41, 43, 45, 270, 272, 274  
20 and 276.

10. The conjugate of claim 1, wherein the therapeutic agent is a toxin, a small organic molecule, a nucleic acid, protein therapeutic agents, a cytokine or a growth factor.

11. The conjugate of claim 1, wherein the therapeutic agent is  
25 an anti-cancer agent.

12. The conjugate of claim 1, wherein the therapeutic agent is an anti-angiogenic agent.

13. The conjugate of claim 1, wherein the therapeutic agent is selected from abrin, ricin A, pseudomonas exotoxin shiga toxin,



-228-

diphtheria toxin, a tumor necrosis factor,  $\alpha$ -interferon,  $\gamma$ -interferon, nerve growth factor, tissue factor and tissue factor variants, FAS-ligand platelet derived growth factor, tissue plasminogen activator, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage  
5 colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), erythropoietin (EPO), nerve growth factor, fibroblast growth factors (FGFs), and epidermal growth factor.

14. The conjugate of claim 1, wherein the therapeutic agent is selected from alkylating agents, toxins, antiproliferative agents, pro-  
10 apoptotic agents, pro-coagulants, cytotoxic nucleosides and tubulin binding agents.

15. The conjugate of claim 1, wherein the therapeutic agent is selected from among the following classes of drugs:

- a) anthracycline family of drugs,
- 15 b) vinca alkaloid drugs,
- c) mitomycins,
- d) bleomycins,
- e) cytotoxic nucleosides,
- f) pteridine family of drugs.
- 20 g) diylenes,
- h) estramustine,
- i) cyclophosphamide,
- j) taxanes,
- k) podophyllotoxins,
- 25 l) maytansanoids,
- m) epothilones, and
- n) combretastatin and analogs,

or pharmaceutically acceptable derivatives thereof.

-229-

16. The conjugate of claim 1, wherein the therapeutic agent is selected from among the following drugs:

- a) doxorubicin,
- b) carminomycin,
- 5 c) daunorubicin,
- d) aminopterin,
- e) methotrexate,
- f) methopterin,
- g) dichloromethotrexate,
- 10 h) mitomycin C,
- i) porfiromycin,
- j) 5-fluorouracil,
- k) 6-mercaptopurine,
- l) cytosine arabinoside,
- 15 m) podophyllotoxin,
- n) etoposide,
- o) etoposide phosphate,
- p) melphalan,
- q) vinblastine,
- 20 r) vincristine,
- s) leurosidine,
- t) vindesine,
- u) estramustine,
- v) cisplatin,
- 25 w) cyclophosphamide,
- x) taxol,
- y) leurositte,
- z) 4-desacetylvinblastine,
- aa) epothilone B,

-230-

- bb) taxotere,
- cc) maytansanol,
- dd) epothilone A, and
- ee) combretastatin and analogs;

5 or a pharmaceutically acceptable derivative thereof.

17. The conjugate of claim 1, further comprising a linker between the therapeutic agent and the peptidic substrate.

18. The conjugate of claim 17, wherein the linker comprises a carbohydrate, peptide, and/or hydrocarbon core.

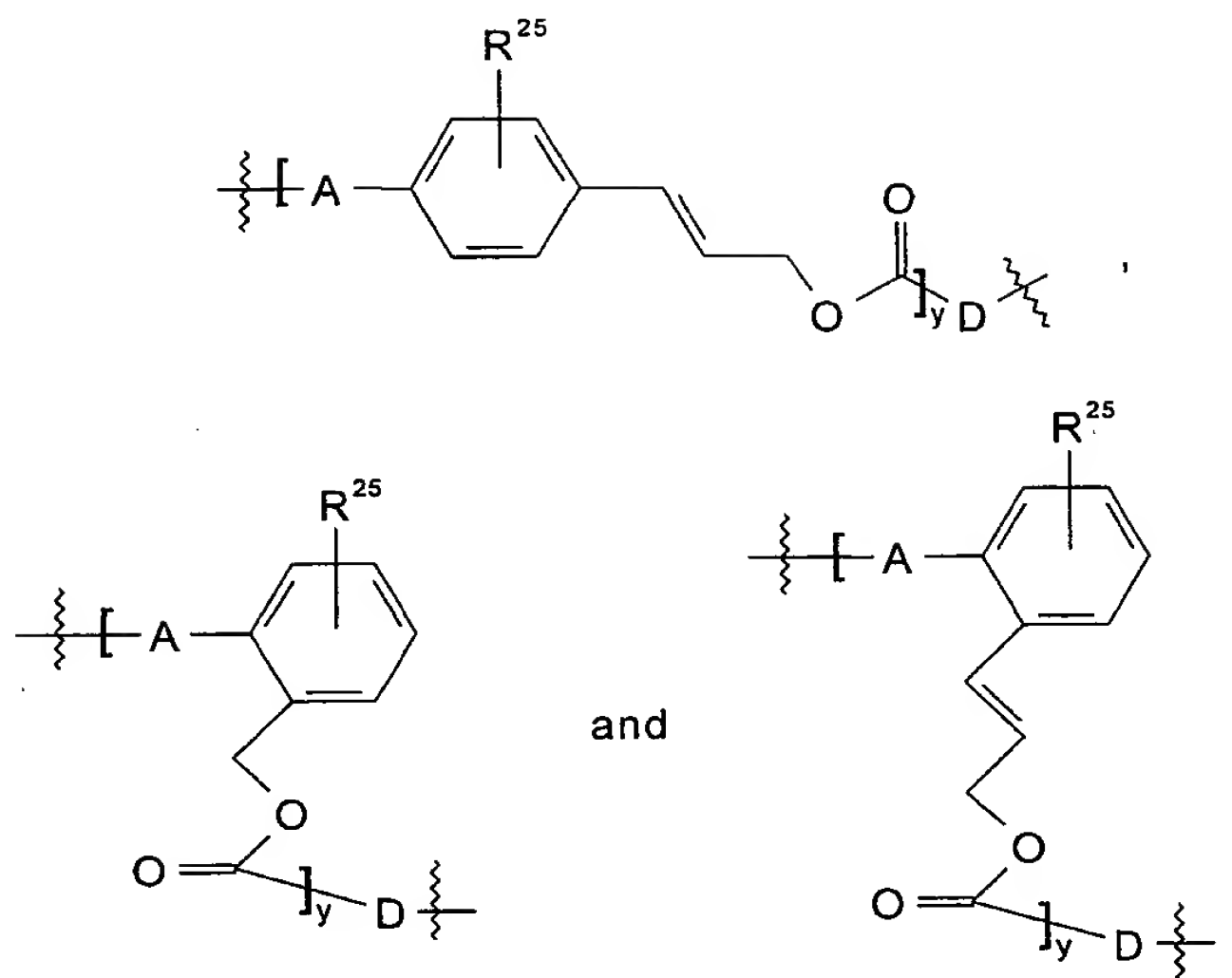
10 19. The conjugate of claim 17, wherein the linker comprises:  
a biscarbonyl alkyl diradical whereby an amine moiety on the therapeutic agent is connected with the linker unit to form an amide bond and the amino terminus of the peptidic substrate is connected with the other end of the linker unit also forming an amide bond; or  
15 a diaminoalkyl diradical linker unit, whereby a carbonyl moiety on the therapeutic agent is covalently attached to one of the amines of the linker unit while the other amine of the linker unit is covalently attached to the C-terminus of the peptidic substrate; or  
is a self-eliminating linker of the following formulae:

20

25

30

-231-



where A is NH or O; D is N(H or alkyl) or O; R<sup>25</sup> is H, alkyl, cycloalkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, such as, for example, 1 to 3, substituents selected from halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2 triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy, haloalkyl and polyhaloalkyl, such as, for example, halo lower alkyl, especially trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that optionally is substituted with 1 or more, such as, for example, 1 to 3, substituents, for example, selected from halo, halo alkyl and alkyl, heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxy carbonyl, aminoimino, alkoxycarbonylamino, aryloxy carbonylamino, aminocarbonyl, alkylaminocarbonyl, dialkylamino-

-232-

carbonyl, arylaminocarbonyl, diarylamino carbonyl, aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkylaminoalkyl, dialkylaminoalkyl, arylaminoalkyl, amino, alkylamino, dialkylamino, arylamino, alkylaryl amino, alkylcarbonylamino, arylcarbonylamino, azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyno, isothiocyno, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl, alkylaminosulfonyl, dialkylaminosulfonyl and arylaminosulfonyl; and y is an integer from 1 to 3.

10           20. The conjugate of claim 17, wherein the linker is a diamine comprising a cyclic alkylene moiety.

          21. The conjugate of claim 17, wherein the diamine contains a bicycloalkylene moiety.

          22. The conjugate of claim 17, wherein the linker selected from  
15 1,4-bis(aminomethyl)cyclohexane, 1,4-bis(aminomethyl)cycloheptane, 1,3-bis(aminomethyl)cyclopentane, 1-amino-4-(aminomethyl)cyclohexane, 1,4-diaminocyclohexane and 1,4-bis(aminomethyl)bicyclo[2.2.2]octane.

          23. The conjugate of claim 17, wherein the linker is a 1, $\omega$ -diaminoalkane.

20           24. The conjugate of claim 17, wherein the linker is a 1,3-diaminopropane.

          25. The conjugate of claim 17, wherein the linker is a 1, $\omega$ -dicarbonylalkane.

          26. The conjugate of claim 25, wherein the linker selected from  
25 oxalic, malonic, succinic, glutaric, adipic and pivalic acids.

          27. The conjugate of claim 1, wherein the peptidic substrate comprises P1 that is any amino acid.

          28. The conjugate of claim 27, wherein P1 is a naturally-occurring amino acid.

-233-

29. The conjugate of claim 27, wherein P1 is an amino acid with an aromatic, branched, or branched aromatic side chain.

30. The conjugate of claim 1, wherein the peptidic substrate comprises P1, where P1 is selected from among Arg, Lys, Tyr, Phe, Trp,  
5 Ala, Val, Ile and Thr.

31. The conjugate of claim 1, wherein:  
the peptidic substrate comprises a P1-P1' bond;  
the P1-P1' bond is the site of cleavage by a cell surface protease;  
P1 is selected from Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;  
10 and  
P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met  
or 6-aminohexanoyl.

32. The conjugate of claim 1, wherein the peptidic substrate comprises P1, wherein P1 is Arg, Lys or an Arg surrogate.

15 33. The conjugate of claim 1, further comprising a P2 residue selected from Phe, Ser, Gly and Ala.

34. The conjugate of claim 1, further comprising a P3 residue selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates.

20 35. The conjugate of claim 1, further comprising a P4 residue selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val.

36. The conjugate of claim 1, further comprising a P5 residue selected from Arg and Arg surrogates.

25 37. The conjugate of claim 1, further comprising a P6 residue selected from Leu, Ile and Val.

38. The conjugate of claim 1, further comprising a P2' residue selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

-234-

39. The conjugate of claim 1, further comprising a P3' residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

40. The conjugate of claim 1, wherein:

5 the peptidic substrate comprises a 5-mer that has the formula:

P4-P3-P2-P1-P1', wherein:

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr

P2 is selected from Phe, Ser, Gly and Ala;

10 P3 is selected from Arg, Lys, Gln, Quat and Arg surrogates;

P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val; and

P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl.

15 41. The conjugate of claim 40, wherein:

the peptidic substrate optionally further comprises one or more of a P5 or P2' amino acid residue, wherein:

P5 is Arg or an Arg surrogate; and

20 P2' is selected from among Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

42. The conjugate of claim 41, wherein:

if the peptidic substrate comprises a P5 amino acid residue, then the peptidic substrate optionally further comprises a P6 amino acid residue selected from Leu, Ile and Val; and

25 if the peptidic substrate comprises a P2' amino acid residue, then the peptidic substrate optionally further comprises a P3' amino acid residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

43. The conjugate of claim 1, wherein:

-235-

the therapeutic agent is conjugated directly or via a linker to the C terminus of the peptidic substrate.

44. The conjugate of claim 1, wherein:

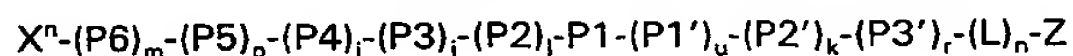
the peptidic substrate comprises a cap at the N-terminus.

5 45. The conjugate of claim 1, wherein the cap is a hydrophilic blocking group.

46. The conjugate of claim 1, wherein the cap is an acyl, sulfonyl or carbamoyl derivative.

47. The conjugate of claim 45, wherein the blocking group is  
10 selected from among hydroxylated alkanoyls, polyhydroxylated alkanoyls, polyethylene glycols, glycosylates, sugars and crown ethers.

48. The conjugate of claim 43 that has formula I:



or a derivative thereof, wherein:

15 Z is a therapeutic agent;

L is a linker;

l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1;  
when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m  
20 are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k and r are selected as follows:

u is 0 or 1; when u is 0, k and r are 0; when u is 1, k is 0 or 1;  
when k is 0, r is 0; when k is 1, r is 0 or 1;

25 n is 0 or 1;

X<sup>n</sup> is hydrogen, or an acyl, sulfonyl or carbamoyl cap;

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;



-236-

P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl;

P2 is selected from Phe, Ser, Gly and Ala;

P3 is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

5 P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

P5 is selected from Arg and Arg surrogates;

P6 is selected from Leu, Ile and Val;

10 P2' is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl; and

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

49. The conjugate of claim 48, wherein P1 is Arg, Lys or an Arg surrogate.

15 50. The conjugate of claim 1, wherein:  
the therapeutic agent is conjugated directly or via a linker to the N terminus of the peptidic substrate.

51. The conjugate of claim 50, wherein:  
the C-terminus of the peptidic substrate is a carboxylic acid or a  
20 carboxamide derivative.

52. The conjugate of claim 50 that has formula II:  
$$Z-(L)_n-(P6)_m-(P5)_p-(P4)_i-(P3)_j-(P2)_l-P1-(P1')_u-(P2')_k-(P3')_r-X^c$$
  
or a derivative thereof, wherein:

Z is a therapeutic agent;

25 L is a linker;

l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1;  
when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m

-237-

are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k and r are selected as follows:

u is 0 or 1; when u is 0, k and r are 0; when u is 1, k is 0 or 1;

5 when k is 0, r is 0; when k is 1, r is 0 or 1;

n is 0 or 1;

X<sup>c</sup>, together with the carbonyl group of the amino acid residue to which it is attached, forms a carboxylic acid or a carboxamide group;

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and  
10 Thr;

P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl;

P2 is selected from Phe, Ser, Gly and Ala;

P3 is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

15 P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

P5 is selected from Arg and Arg surrogates;

P6 is selected from Leu, Ile and Val;

P2' is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal,  
20 Aib, Abu, Met and 6-aminohexanoyl; and

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

53. The conjugate of claim 52, wherein P1 is Arg, Lys or an Arg surrogate.

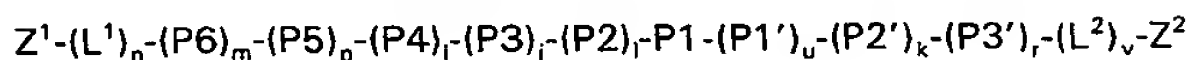
25 54. The conjugate of claim 1, wherein a first therapeutic agent is attached, optionally via a first linker, to the N-terminus of the peptidic substrate; and

a second therapeutic agent, which are the same or different from the first therapeutic agent, is attached, optionally via a second linker,

-238-

which are the same or different from the first linker, to the C-terminus of the peptidic substrate.

55. The conjugate of claim 54 that has formula III:



5 or a derivative thereof, wherein:

$Z^1$  and  $Z^2$  are each therapeutic agents and are the same or different;

$L^1$  and  $L^2$  are each linkers and are the same or different;

$l, j, i, p$  and  $m$  are selected as follows:

10  $l$  is 0 or 1; when  $l$  is 0,  $j, i, p$  and  $m$  are 0; when  $l$  is 1,  $j$  is 0 or 1; when  $j$  is 0,  $i, p$  and  $m$  are 0; when  $j$  is 1,  $i$  is 0 or 1; when  $i$  is 0,  $p$  and  $m$  are 0; when  $i$  is 1,  $p$  is 0 or 1; when  $p$  is 0,  $m$  is 0; when  $p$  is 1,  $m$  is 0 or 1;

$u, k$  and  $r$  are selected as follows:

15  $u$  is 0 or 1; when  $u$  is 0,  $k$  and  $r$  are 0; when  $u$  is 1,  $k$  is 0 or 1; when  $k$  is 0,  $r$  is 0; when  $k$  is 1,  $r$  is 0 or 1;

$n$  and  $v$  are each independently 0 or 1;

$P1$  is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;

20  $P1'$  is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl;

$P2$  is selected from Phe, Ser, Gly and Ala;

$P3$  is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

25  $P4$  is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

$P5$  is selected from Arg and Arg surrogates;

$P6$  is selected from Leu, Ile and Val;

$P2'$  is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl; and

-239-

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

56. The conjugate of claim 55, wherein P1 is Arg, Lys or an Arg surrogate.

- 5        57. The conjugate of any of claims 1-57, selected from:  
Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 46);  
Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 47);
- 10    Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 48);  
Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 49);  
Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 50);
- 15    Ac-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 51);  
Ac-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 52);  
Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 53);  
Ac-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 54);  
Ac-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 55);
- 20    Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 56);  
Ac-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 57);  
Ac-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 58);  
Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 59);  
Ac-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 60);
- 25    Ac-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 61);  
Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 62);  
Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 63);

-240-

- Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 64);
- Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 65);
- 5 Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 66);  
 Ac-Arg-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 67);  
 Ac-Pro-Arg-Phe-Lys-Ile-Ile-(therapeutic agent) (SEQ ID NO: 68);  
 Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 69);  
 Ac-Arg-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 70);
- 10 Ac-Ser-Lys-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 71);  
 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 72);  
 Ac-Arg-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 73);  
 Ac-Pro-Arg-Phe-Arg-Ile-Ile-(therapeutic agent) (SEQ ID NO: 74);  
 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 75);
- 15 Ac-Arg-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 76); and  
 Ac-Ser-Arg-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 77)  
 pyroGlu-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 78);  
 CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 79);  
 N-p-tosyl-Gly-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 80);
- 20 Benzoyl-Val-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 81);  
 CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 82);  
 N- $\alpha$ -Z-D-Arg-Gly-Arg-Ala-Ala-(therapeutic agent) in which Z is  
 benzyloxycarbonyl (SEQ ID NO: 83);  
 pyroGlu-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 84);
- 25 H-D-Ile-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 85);  
 Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ala-Ala-(therapeutic agent) in which Cbo is  
 carbobenzoxo (SEQ ID NO: 86);  
 H-D-Pro-Phe-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 87);  
 H-D-Val-Leu-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 88);

-241-

- Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ala-Ala-(therapeutic agent) in which Bz is benzoyl (SEQ ID NO: 89);
- Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 90);
- Bz-Pro-Phe-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 91);
- 5 H-D-Phe-Pip-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 92);
- H-D-Val-Leu-Lys-Ala-Ala-(therapeutic agent) (SEQ ID NO: 93);
- H-D-Nle-HHT-Lys-Ala-Ala-(therapeutic agent) (SEQ ID NO: 94);
- Pyr-Arg-Thr-Lys-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 95);
- H-Arg-Gln-Arg-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 96);
- 10 Boc-Gln-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 97);
- Z-Arg-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 98);
- H-D-HHT-Ala-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 99);
- H-D-CHT-Gly-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 100);
- MeSO<sub>2</sub>-dPhe-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 101);
- 15  $\delta$ -Z-D-Lys-Pro-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 102);
- CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 103);
- Ac-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 104);
- Ac-Arg-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 105);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-Ala-(therapeutic agent) (SEQ ID NO: 106);
- 20 Ac-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 107);
- Ac-Arg-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 108);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Gly-Gly-(therapeutic agent) (SEQ ID NO: 109);
- 25 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-(therapeutic agent) (SEQ ID NO: 110);
- Ac-Arg-Arg-Gln-Ser-Arg-Ile-(therapeutic agent) (SEQ ID NO: 111);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ala-Ile-(therapeutic agent) (SEQ ID NO: 112);
- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 113);

-242-

- Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 114);
- Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 115);
- 5 Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 116);
- Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 117);
- Ac-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 118);
- 10 Ac-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 119);
- Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 120);
- Ac-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 121);
- Ac-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 122);
- 15 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 123);
- Ac-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 124);
- Ac-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 125);
- Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 126);
- 20 Ac-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 127);
- Ac-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 128);
- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 129);
- 25 Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 130);
- Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 131);

-243-

- Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 132);
- Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 133);
- 5 Ac-Arg-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 134);  
 Ac-Pro-Arg-Phe-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 135);  
 Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 136);
- Ac-Arg-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 137);
- 10 Ac-Ser-Lys-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 138);  
 Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 139);
- Ac-Arg-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 140);  
 Ac-Pro-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 141);
- 15 Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 142);
- Ac-Arg-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 143);  
 Ac-Ser-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 144);  
 pyroGlu-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 145);
- 20 CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 146);  
 N-p-tosyl-Gly-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 147);  
 Benzoyl-Val-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 148);  
 CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 149);  
 N- $\alpha$ -Z-D-Arg-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 150) (Z =
- 25 benzyloxycarbonyl);
- pyroGlu-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 151);  
 H-D-Ile-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 152);  
 Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 153) (Cbo = carbobenzoxy);



-244-

- H-D-Pro-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 154);  
H-D-Val-Leu-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 155);  
Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 156)  
(Bz = benzoyl);
- 5 Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 157);  
Benzoyl-Pro-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 158);  
H-D-Phe-Pip-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 159);  
H-D-Val-Leu-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 160);  
H-D-Nle-HHT-Lys-Ser-Leu-(therapeutic agent) (SEQ ID NO: 161);
- 10 Pyr-Arg-Thr-Lys-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 162);  
H-Arg-Gln-Arg-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 163);  
Boc-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 164);  
Z-Arg-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 165);  
H-D-HHT-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 166);
- 15 H-D-CHT-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 167);  
MeSO<sub>2</sub>-dPhe-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 168);  
 $\delta$ -Z-D-Lys-Pro-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 169);  
CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 170);  
Ac-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 171);
- 20 Ac-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 172);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 173);  
Ac-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 174);  
Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 175);
- 25 Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 176);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 177);  
Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 178);

-245-

- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 179);
- Ac-Arg-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 180);
- Ac-Arg-Gln-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 181);
- 5 Ac-Arg-Gln-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 182);
- Ac-Arg-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 183);
- Ac-Arg-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 184);
- Ac-Arg-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 185);
- Ac-Arg-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 186);
- 10 Ac-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 187);
- Ac-Gln-Gly-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 188);
- Ac-Gln-Ala-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 189);
- Ac-Gln-Phe-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 190).
- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 191);
- 15 Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 192);
- Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 193);
- 20 Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 194);
- Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 195);
- Ac-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 196);
- 25 Ac-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 197);
- Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 198);

-246-

- Ac-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 199);
- Ac-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 200);
- Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 5 201);
- Ac-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 202);
- Ac-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 203);
- Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 10 204);
- Ac-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 205);
- Ac-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 206);
- Ac-Leu-Arg-Ala-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 15 207);
- Ac-Leu-Arg-Ala-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 208);
- Ac-Leu-Arg-Ser-Quat-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 209);
- 20 Ac-Leu-Arg-Ser-Quat-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 210);
- Ac-Leu-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 211);
- Ac-Arg-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 25 212);
- Ac-Pro-Arg-Phe-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 213);
- Ac-Leu-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 214);

-247-

- Ac-Arg-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 215);
- Ac-Ser-Lys-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 216);
- Ac-Leu-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 217);
- 5 Ac-Arg-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 218);
- Ac-Pro-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 219);
- Ac-Leu-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 220);
- 10 Ac-Arg-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 221);
- Ac-Ser-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 222);
- pyroGlu-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 223);
- 15 CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 224);
- N-p-tosyl-Gly-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 225);
- Benzoyl-Val-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 226);
- CH<sub>3</sub>SO<sub>2</sub>-D-HHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 227);
- 20 N- $\alpha$ -Z-D-Arg-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 228) (Z = benzyloxycarbonyl);
- pyroGlu-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 229);
- H-D-Ile-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 230);
- 25 Cbo-L-( $\gamma$ )Glu( $\alpha$ -t-BuO)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 231) (Cbo = carbobenzoxy);
- H-D-Pro-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 232);
- H-D-Val-Leu-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 233);

-248-

- Bz-Ile-Glu( $\gamma$ -OH)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 234) (Bz = benzoyl);
- Bz-Ile-Glu( $\gamma$ -OMe)-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 235);
- 5 Benzoyl-Pro-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 236);
- H-D-Phe-Pip-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 237);
- H-D-Val-Leu-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 238);
- H-D-Nle-HHT-Lys-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 239);
- Pyr-Arg-Thr-Lys-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 240);
- 10 H-Arg-Gln-Arg-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 241);
- Boc-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 242);
- Z-Arg-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 243);
- H-D-HHT-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 244);
- H-D-CHT-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 245);
- 15 MeSO<sub>2</sub>-dPhe-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 246);
- $\delta$ -Z-D-Lys-Pro-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 247);
- CH<sub>3</sub>SO<sub>2</sub>-D-CHA-But-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 248);
- Ac-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 249);
- 20 Ac-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 250);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 251);
- Ac-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 252);
- 25 Ac-Arg-Arg-Gln-Ser-Arg-Leu-(therapeutic agent) (SEQ ID NO: 253);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 254);
- Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 255);

-249-

- Ac-Arg-Arg-Gln-Ser-Arg-Ser-Leu-(therapeutic agent) (SEQ ID NO: 256);  
Ac-Leu-Arg-Arg-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 257);  
Ac-Arg-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 258);  
5 Ac-Arg-Gln-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 259);  
Ac-Arg-Gln-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 260);  
Ac-Arg-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 261);  
Ac-Arg-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 262);  
Ac-Arg-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 263);  
10 Ac-Arg-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 264);  
Ac-Gln-Ser-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 265);  
Ac-Gln-Gly-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 266);  
Ac-Gln-Ala-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 267); and  
Ac-Gln-Phe-Arg-Ser-Ser-Leu-(therapeutic agent) (SEQ ID NO: 268).
- 15 58. The conjugate of claim 35, wherein P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Phe and Val.
59. The conjugate of claim 35, wherein:  
P2, P3 and/or P4 is/are selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, GLu, Phe and Val.
- 20 60. The conjugate of claim 35, wherein:  
P2, P3 and/or P4 is/are selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Tyr, Glu, Leu Phe and Val; and  
P1 is any amino acid.
61. The conjugate of claim 60, wherein P1 is a naturally-  
25 occurring amino acid.
62. The conjugate of claim 60, wherein P1 is an amino acid with an aromatic, branched, or branched aromatic side chain.
63. The conjugate of claim 60, wherein P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr.

-250-

64. The conjugate of claim 60, wherein P1 is Arg, Lys or an Arg surrogate.

65. The conjugate of claim 1, wherein the protease is located at the cell surface by virtue of a specific binding interaction with a receptor  
5 therefor.

66. The conjugate of claim 65, wherein the cell surface protease is urokinase plasminogen activator (u-PA) bound to urokinase plasminogen activator receptor (u-PAR).

67. The conjugate of claim 1, that comprises a peptidic  
10 substrate of the formula P6-P5-P4-P3-P2-P1-P1'-P2'-P3', wherein each of P1, P2, P3, P4, P5, P6, P1' and P2' are selected from residues set forth in Figures 1 and 2, and P6, P5, P4, P2' and P3' are optional.

68. The conjugate of claim 67, wherein:  
P6 is optional and is selected from L, V, R;  
15 P5 is optional and is selected from R, I, L;  
P4 is optional and is selected from G, C, V;  
P3 is selected from S, dS, P, A or G;  
P2 is selected from A or G;  
P1 is R;  
20 P1' is S, V, M or nL;  
P2' is optional and is selected S, L, A or V; and  
P3' is optional and is L.

69. A conjugate selected from among those set forth in Figures 1-5, wherein the therapeutic agent doxorubicin (Dox) or taxol (Tax)  
25 optionally is replaced with any therapeutic agent.

70. The conjugate of claim 65, wherein the therapeutic agent is a toxin, a small organic molecule, a nucleic acid, protein therapeutic agents, a cytokine or a growth factor.

-251-

71. The conjugate of claim 65, wherein the therapeutic agent is an anti-cancer agent.

72. The conjugate of claim 65, wherein the therapeutic agent is an anti-angiogenic agent.

5        73. The conjugate of claim 65, wherein the therapeutic agent is selected from abrin, ricin A, pseudomonas exotoxin shiga toxin, diphtheria toxin, a tumor necrosis factor,  $\alpha$ -interferon,  $\gamma$ -interferon, nerve growth factor, tissue factor and tissue factor variants, FAS-ligand platelet derived growth factor, tissue plasminogen activator, interleukin-1  
10 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), erythropoietin (EPO), nerve growth factor, fibroblast growth factors (FGFs), and epidermal growth factor.

74. The conjugate of claim 65, wherein the therapeutic agent is  
15 selected from alkylating agents, toxins, antiproliferative agents, pro-apoptotic agents, pro-coagulants, cytotoxic nucleosides and tubulin binding agents.

75. The conjugate of claim 65, wherein the therapeutic agent is selected from among the following classes of drugs:

- 20            a) anthracycline family of drugs,  
              b) vinca alkaloid drugs,  
              c) mitomycins,  
              d) bleomycins,  
              e) cytotoxic nucleosides,  
25            f) pteridine family of drugs.  
              g) diynenes,  
              h) estramustine,  
              i) cyclophosphamide,  
              j) taxanes,



-252-

- k) podophyllotoxins,
- l) maytansanoids,
- m) epothilones, and
- n) combretastatin and analogs,

5 or pharmaceutically acceptable derivatives thereof.

76. The conjugate of claim 65, wherein the therapeutic agent is selected from among the following drugs:

- a) doxorubicin,
- b) carminomycin,
- 10 c) daunorubicin,
- d) aminopterin,
- e) methotrexate,
- f) methopterin,
- g) dichloromethotrexate,
- 15 h) mitomycin C,
- i) porfiromycin,
- j) 5-fluorouracil,
- k) 6-mercaptopurine,
- l) cytosine arabinoside,
- 20 m) podophyllotoxin,
- n) etoposide,
- o) etoposide phosphate,
- p) melphalan,
- q) vinblastine,
- 25 r) vincristine,
- s) leurosidine,
- t) vindesine,
- u) estramustine,
- v) cisplatin,

-253-

- w) cyclophosphamide,
- x) taxol,
- y) leurositte,
- z) 4-desacetylvinblastine,
- 5 aa) epothilone B,
- bb) taxotere,
- cc) maytansanol,
- dd) epothilone A, and
- ee) combretastatin and analogs;

10 or a pharmaceutically acceptable derivative thereof.

77. The conjugate of claim 65, further comprising a linker between the therapeutic agent and the peptidic substrate.

78. The conjugate of claim 65, wherein the linker comprises a carbohydrate, peptide, and/or hydrocarbon core.

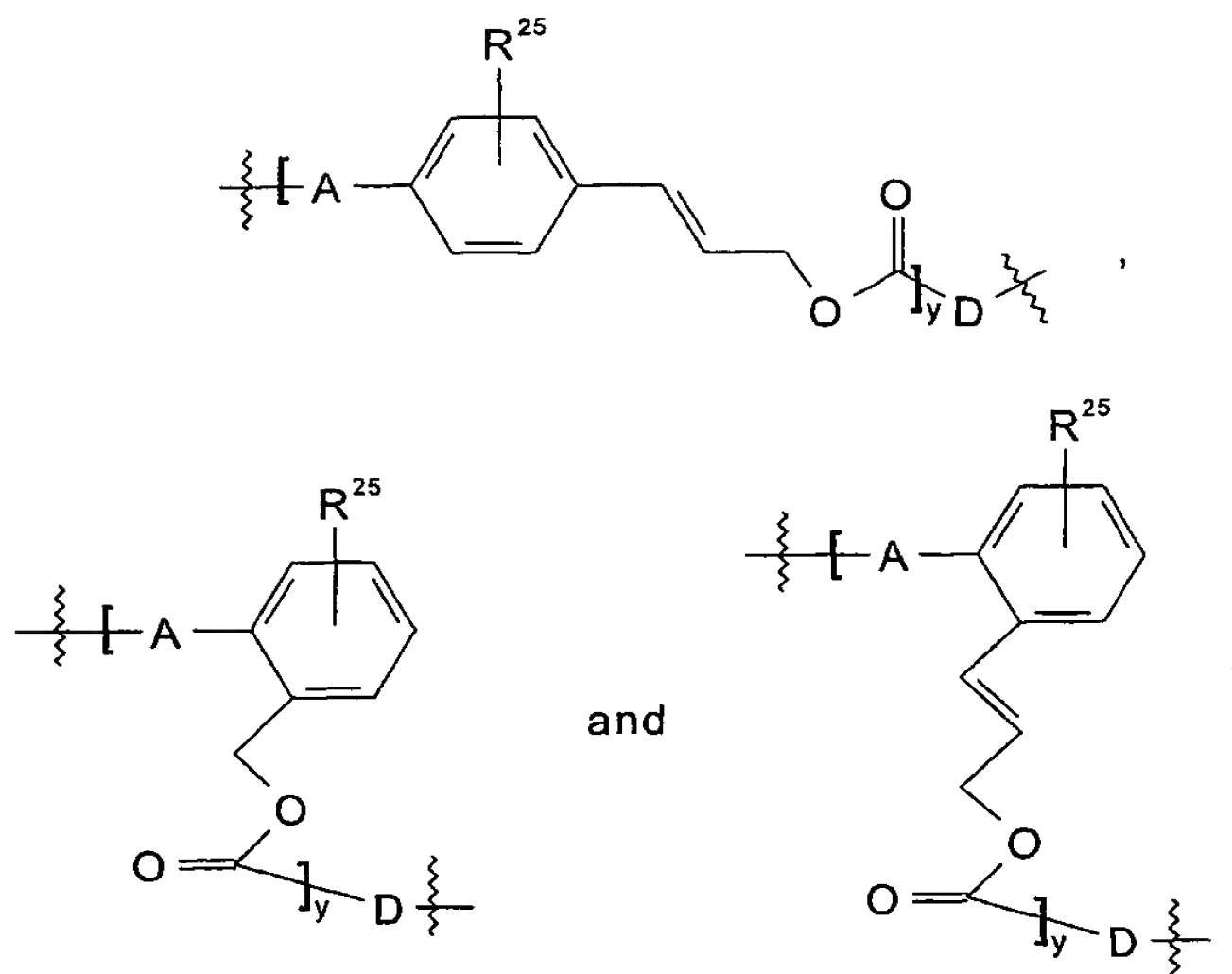
15 79. The conjugate of claim 77, wherein the linker comprises:  
a biscarbonyl alkyl diradical whereby an amine moiety on the therapeutic agent is connected with the linker unit to form an amide bond and the amino terminus of the peptidic substrate is connected with the other end of the linker unit also forming an amide bond; or

20 a diaminoalkyl diradical linker unit, whereby a carbonyl moiety on the therapeutic agent is covalently attached to one of the amines of the linker unit while the other amine of the linker unit is covalently attached to the C-terminus of the peptidic substrate; or

is a self-eliminating linker of the following formulae:

25

-254-



where A is NH or O; D is N(H or alkyl) or O; R<sup>25</sup> is H, alkyl, cycloalkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, such as 1 to 3, substituents selected from, for example, halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2 triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy, haloalkyl and polyhaloalkyl, such as, for example, halo lower alkyl, including trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that optionally is substituted with 1 or more, such as, for example, 1 to 3, substituents selected from, for example, halo, halo alkyl and alkyl, heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxycarbonyl, aminoimino, alkoxycarbonylamino, aryloxycarbonylamino, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, arylaminocarbonyl,

-255-

diarylaminocarbonyl, aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkyl-aminoalkyl, dialkylaminoalkyl, arylaminoalkyl, amino, alkylamino, dialkyl-amino, arylamino, alkylarylamino, alkylcarbonylamino, arylcarbonylamino, 5 azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyno, isothiocyno, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl, alkylaminosulfonyl, dialkylaminosulfonyl and arylamino-sulfonyl; and  $y$  is an integer from 1 to 3.

80. The conjugate of claim 77, wherein the linker is a diamine  
10 comprising a cyclic alkylene moiety.

81. The conjugate of claim 77, wherein the diamine contains a bicycloalkylene moiety.

82. The conjugate of claim 77, wherein the linker selected from  
1,4-bis(aminomethyl)cyclohexane, 1,4-bis(aminomethyl)cycloheptane,  
15 1,3-bis(aminomethyl)cyclopentane, 1-amino-4-(aminomethyl)cyclohexane,  
1,4-diaminocyclohexane and 1,4-bis(aminomethyl)bicyclo[2.2.2]octane.

83. The conjugate of claim 77, wherein the linker is a 1, $\omega$ -diaminoalkane.

84. The conjugate of claim 77, wherein the linker is a  
20 1,3-diaminopropane.

85. The conjugate of claim 77, wherein the linker is a 1, $\omega$ -dicarbonylalkane.

86. The conjugate of claim 77, wherein the linker is selected from oxalic, malonic, succinic, glutaric, adipic and pivalic acids.

25 87. The conjugate of any of claims 1-30, wherein:  
the peptidic substrate comprises a P1-P1' bond;  
the P1-P1' bond is the site of cleavage by a cell surface protease;  
P1 is selected from Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;  
and

-256-

P1' is Gly, Ser, hSer, Thr, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl.

88. The conjugate of any of claims 1-32, further comprising a P2 residue selected from Phe, Ser, Gly, Ala, Ser(OMe), hSer, 1-methylHis, 3-methylHis, His, nVal, nLeu, Abu, (hS)Gly, Thr, Aib, CHA and Tyr.

89. The conjugate of any of claims 1-33, further comprising a P3 residue selected from Arg, Lys, Gln, Quat, Arg surrogates, Ser, Thr, hSer, dSer, Pro, (hS)Gly, Tyr, 4,4-dimethylThr, Asn, Met(O<sub>2</sub>), Quat<sup>2</sup>, Quat<sup>3</sup>, Quat<sup>4</sup> and Quat<sup>5</sup>.

90. The conjugate of any of claims 1-34, further comprising a P4 residue selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe, Val, N,N-dimethylGly,  $\beta$ -Ala, Cys(Me), Gln, t-butylGly and nVal.

91. The conjugate of any of claims 1-35, further comprising a P5 residue selected from Ile, Arg and Arg surrogates.

92. The conjugate of any of claims 1-36, further comprising a P6 residue selected from Val, Leu, Ile and Val.

93. The conjugate of any of claims 1-37, further comprising a P2' residue selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, hCHA, CHA, hexylGly, allylGly and Phe.

94. The conjugate of any of claims 1-38, further comprising a P3' residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly.

95. The conjugate of any of claims 1-39, further comprising a P4' residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly.

96. The conjugate of any of claims 1-39, wherein P4' is Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

97. The conjugate of any of claims 1-39, wherein P4' is Leu.

98. The conjugate of any of claims 1-39, wherein:

-257-

the peptidic substrate comprises a 5-mer that has the formula:

P4-P3-P2-P1-P1', wherein:

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;

5 P2 is selected from Phe, Ser, Gly, Ala, Ser(OMe), hSer, 1-methylHis, 3-methylHis, His, nVal, nLeu, Abu, (hS)Gly, Thr, Aib, CHA and Tyr;

P3 is selected from Arg, Lys, Gln, Quat, Arg surrogates, Ser, Thr, hSer, dSer, Pro, (hS)Gly, Tyr, 4,4-dimethylThr, Asn, Met(O<sub>2</sub>), Quat<sup>2</sup>,  
10 Quat<sup>3</sup>, Quat<sup>4</sup> and Quat<sup>5</sup>;

P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe, Val, N,N-dimethylGly,  $\beta$ -Ala, Cys(Me), Gln, t-butylGly and nVal; and

P1' is Gly, Ser, hSer, Thr, Ala, Leu, Ile, d-Ile, nLeu, Val,  
15 nVal, Aib, Abu, Met or 6-aminohexanoyl.

99. The conjugate of claim 40, wherein:

the peptidic substrate optionally further comprises one or more of a P5 or P2' amino acid residue, wherein:

P5 is Ile, Arg or an Arg surrogate; and

20 P2' is selected from among Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, hCHA, CHA, hexylGly, allylGly and Phe.

100. The conjugate of claim 41, wherein:

if the peptidic substrate comprises a P5 amino acid residue, then

25 the peptidic substrate optionally further comprises a P6 amino acid residue selected from Arg, Leu, Ile and Val; and

if the peptidic substrate comprises a P2' amino acid residue, then the peptidic substrate optionally further comprises a P3' amino acid

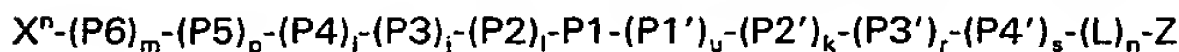
-258-

residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly; and

if the peptidic substrate comprises a P3' amino acid residue, then the peptidic substrate optionally further comprises a P4' amino acid

5 residue selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met, 6-aminohexanoyl, CHA and allylGly.

101. The conjugate of any of claims 43-47 that has formula IV:



or a derivative thereof, wherein:

10 Z is a therapeutic agent;

L is a linker;

l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1;  
when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m  
15 are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or 1;

u, k, r and s are selected as follows:

u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1;  
when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0;  
20 when r is 1, s is 0 or 1;

n is 0 or 1;

X<sup>n</sup> is hydrogen, or an acyl, sulfonyl or carbamoyl cap;

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;

25 P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl;

P2 is selected from Phe, Ser, Gly and Ala;

P3 is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

-259-

P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

P5 is selected from Arg and Arg surrogates;

P6 is selected from Leu, Ile and Val;

5 P2' is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl;

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl; and

10 P4' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

102. The conjugate of claim 50 or claim 51 that has formula V:  

$$Z-(L)_n-(P6)_m-(P5)_p-(P4)_i-(P3)_j-(P2)_l-P1-(P1')_u-(P2')_k-(P3')_r-(P4')_s-X^c$$
  
 or a derivative thereof, wherein:

Z is a therapeutic agent;

15 L is a linker;

l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or

20 1;

u, k, r and s are selected as follows:

u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1; when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0; when r is 1, s is 0 or 1;

25 n is 0 or 1;

X<sup>c</sup>, together with the carbonyl group of the amino acid residue to which it is attached, forms a carboxylic acid or a carboxamide group;

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;



-260-

P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met or 6-aminohexanoyl;

P2 is selected from Phe, Ser, Gly and Ala;

P3 is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

5 P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

P5 is selected from Arg and Arg surrogates;

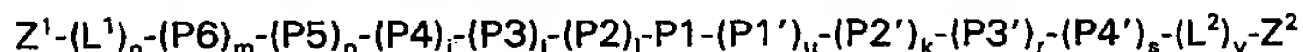
P6 is selected from Leu, Ile and Val;

10 P2' is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl;

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl; and

P4' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

15 103. The conjugate of claim 54 that has formula VI:



or a derivative thereof, wherein:

Z<sup>1</sup> and Z<sup>2</sup> are each therapeutic agents and are the same or different;

20 L<sup>1</sup> and L<sup>2</sup> are each linkers and are the same or different;

l, j, i, p and m are selected as follows:

l is 0 or 1; when l is 0, j, i, p and m are 0; when l is 1, j is 0 or 1; when j is 0, i, p and m are 0; when j is 1, i is 0 or 1; when i is 0, p and m are 0; when i is 1, p is 0 or 1; when p is 0, m is 0; when p is 1, m is 0 or

25 1;

u, k, r and s are selected as follows:

u is 0 or 1; when u is 0, k, r and s are 0; when u is 1, k is 0 or 1; when k is 0, r and s are 0; when k is 1, r is 0 or 1; when r is 0, s is 0; when r is 1, s is 0 or 1;

-261-

n and v are each independently 0 or 1;

P1 is selected from among Arg, Lys, Tyr, Phe, Trp, Ala, Val, Ile and Thr;

P1' is Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met  
5 or 6-aminohexanoyl;

P2 is selected from Phe, Ser, Gly and Ala;

P3 is selected from Arg, Lys, Gln, Ser, Quat and Arg surrogates;

P4 is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe and Val;

10 P5 is selected from Arg and Arg surrogates;

P6 is selected from Leu, Ile and Val;

P2' is selected from Gly, Ser, Ala, Leu, Ile, d-Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl;

P3' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib,  
15 Abu, Met and 6-aminohexanoyl; and

P4' is selected from Gly, Ser, Ala, Leu, Ile, nLeu, Val, nVal, Aib, Abu, Met and 6-aminohexanoyl.

104. The conjugate of any of claims 1-49, selected from:

Ac-R-Q-G-R-S-L-(therapeutic agent) (SEQ ID NO: 491);

20 Ac-R-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 492);

Ac-R-Q-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 493);

Ac-R-Q-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 494);

Ac-R-Q-G-R-S-F-(therapeutic agent) (SEQ ID NO: 495);

Ac-R-Q-G-R-A-L-(therapeutic agent) (SEQ ID NO: 496);

25 Ac-R-Q-G-R-A-L-(therapeutic agent) (SEQ ID NO: 497);

Ac-R-Q-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 498);

Ac-R-Q-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 499);

Ac-R-Q-G-R-A-nV-(therapeutic agent) (SEQ ID NO: 500);

Ac-R-Q-G-R-A-Cha-(therapeutic agent) (SEQ ID NO: 501);

-262-

- Ac-R-Q-G-R-A-F-(therapeutic agent) (SEQ ID NO: 502);  
Ac-R-N-G-R-S-L-(therapeutic agent) (SEQ ID NO: 503);  
Ac-R-N-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 504);  
Ac-R-Q-A-R-S-L-(therapeutic agent) (SEQ ID NO: 505);  
5 Ac-R-Q-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 506);  
Ac-R-Q-A-R-S-nV-(therapeutic agent) (SEQ ID NO: 507);  
Ac-R-Q-A-A-S-Cha-(therapeutic agent) (SEQ ID NO: 508);  
Ac-R-Q-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 509);  
Ac-R-Q-A-R-T-nL-(therapeutic agent) (SEQ ID NO: 510);  
10 Ac-R-Q-A-R-A-L-(therapeutic agent) (SEQ ID NO: 511);  
Ac-R-Q-A-R-A-nL-(therapeutic agent) (SEQ ID NO: 512);  
Ac-R-Q-A-R-A-nV-(therapeutic agent) (SEQ ID NO: 513);  
Ac-R-Q-A-R-A-Cha-(therapeutic agent) (SEQ ID NO: 514);  
Ac-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 515);  
15 Ac-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 516);  
Ac-R-Q-S-R-A-nL-(therapeutic agent) (SEQ ID NO: 517);  
Ac-R-Q-S-R-A-L-(therapeutic agent) (SEQ ID NO: 518);  
Ac-R-Q-S-R-A-nV-(therapeutic agent) (SEQ ID NO: 519);  
Ac-R-Q-S-R-A-Cha-(therapeutic agent) (SEQ ID NO: 520);  
20 Ac-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 521);  
Ac-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 522);  
Ac-R-Q-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 523);  
Ac-R-Q-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 524);  
Ac-R-Q-S-R-S-nV-(therapeutic agent) (SEQ ID NO: 525);  
25 Ac-R-Q-S-R-S-allylG-(therapeutic agent) (SEQ ID NO: 526);  
Ac-R-Q-S-R-S-Cha-(therapeutic agent) (SEQ ID NO: 527);  
Ac-R-Q-S-R-T-nL-(therapeutic agent) (SEQ ID NO: 528);  
Ac-R-Q-T-R-S-S-L-(therapeutic agent) (SEQ ID NO: 529);  
Ac-R-Q-T-R-S-L-(therapeutic agent) (SEQ ID NO: 530);

-263-

- Ac-R-N-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 531);  
Ac-R-Q-F-R-S-L-(therapeutic agent) (SEQ ID NO: 532);  
Ac-R-Q-F-R-S-nL-(therapeutic agent) (SEQ ID NO: 534);  
Ac-R-Q-F-R-S-nV-(therapeutic agent) (SEQ ID NO: 535);  
5 Ac-R-Q-F-R-S-nL-(therapeutic agent) (SEQ ID NO: 536);  
Ac-R-Q-F-R-S-Cha-(therapeutic agent) (SEQ ID NO: 537);  
Ac-R-Q-F-R-A-L-(therapeutic agent) (SEQ ID NO: 538);  
Ac-R-Q-F-R-A-nL-(therapeutic agent) (SEQ ID NO: 539);  
Ac-R-Q-F-R-A-nV-(therapeutic agent) (SEQ ID NO: 540);  
10 Ac-R-Q-F-R-A-Cha-(therapeutic agent) (SEQ ID NO: 541);  
Ac-Q-S-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 542);  
MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 483);  
MeOCO-Quat3-G-R-S-L-(therapeutic agent) (SEQ ID NO: 484);  
MeOCO-Quat-G-R-S-L-(therapeutic agent) (SEQ ID NO: 485);  
15 MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 486);  
MeOCO-Quat5-G-R-S-L-(therapeutic agent) (SEQ ID NO: 487);  
MeOCO-Quat2-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 488);  
MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 489);  
MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 490);  
20 Ac-Q-G-R-S-L-(therapeutic agent) (SEQ ID NO: 445);  
Ac-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 446);  
Ac-Q-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 447);  
Ac-N-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 448);  
Ac-Q-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 449);  
25 Ac-Q-G-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 450);  
Ac-Q-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 451);  
Ac-Q-G-R-S-S-allylG-(therapeutic agent) (SEQ ID NO: 452);  
Ac-Q-G-R-S-S-allylG-(therapeutic agent) (SEQ ID NO: 453);  
Ac-Q-A-R-S-L-(therapeutic agent) (SEQ ID NO: 454);

-264-

- Ac-Q-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 455);  
Ac-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 456);  
Ac-Q-S-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 457);  
Ac-Q-S-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 458);  
5 Ac-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 459);  
Ac-Q-T-R-S-S-L-(therapeutic agent) (SEQ ID NO: 460);  
Ac-Q-Aib-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 461);  
Ac-Q-Aib -R-S-S-L-(therapeutic agent) (SEQ ID NO: 462);  
Ac-Q-Abu-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 463);  
10 Ac-Q-Abu-R-S-S-L-(therapeutic agent) (SEQ ID NO: 464);  
Ac-Q-Cha-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 465);  
Ac-Q-F-R-S-L-(therapeutic agent) (SEQ ID NO: 466);  
Ac-Q-F-R-S-S-L-(therapeutic agent) (SEQ ID NO: 467);  
Ac-Q-Y-R-S-S-L-(therapeutic agent) (SEQ ID NO: 468);  
15 Ac-R-G-R-S-L-(therapeutic agent) (SEQ ID NO: 469);  
Ac-R-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 470);  
Ac-R-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 471);  
Ac-R-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 472);  
Ac-R-A-R-S-L-(therapeutic agent) (SEQ ID NO: 473);  
20 Ac-R-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 474);  
Ac-R-S-R-S-L-(therapeutic agent) (SEQ ID NO: 475);  
Ac-R-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 476);  
Ac-R-S-R-S-Cha-(therapeutic agent) (SEQ ID NO: 477);  
Ac-R-S-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 478);  
25 Ac-R-F-R-S-L-(therapeutic agent) (SEQ ID NO: 479);  
Ac-R-F-R-S-Cha-(therapeutic agent) (SEQ ID NO: 480);  
Ac-Y-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 481);  
Ac-M(O2)-S-R-S-L-(therapeutic agent) (SEQ ID NO: 482);  
Ac-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 105);

-265-

- Ac-R-R-Q-S-R-I-(therapeutic agent) (SEQ ID NO: 610);  
Ac-R-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 543);  
Ac-R-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 544);  
Ac-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 545);  
5 Ac-R-G-S-G-R--S-nL-(therapeutic agent) (SEQ ID NO: 546);  
Ac-R-G-S-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 547);  
Ac-R-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 548);  
Ac-I-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 549);  
Ac-R-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 108);  
10 Ac-R-R-Q-S-R-I-(therapeutic agent) (SEQ ID NO: 111);  
Ac-L-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 106);  
Ac-L-R-R-Q-S-R-G-G-(therapeutic agent) (SEQ ID NO: 109);  
Ac-L-R-R-Q-S-R-A-(therapeutic agent) (SEQ ID NO: 110);  
Ac-L-R-R-Q-S-R-A-I-(therapeutic agent) (SEQ ID NO: 112);  
15 Ac-L-R-R-Q-S-R-A-I-(therapeutic agent) (SEQ ID NO: 611);  
Ac-L-R-R-Q-S-R-S-S-L-(therapeutic agent) (SEQ ID NO: 550);  
Ac-L-R-R-Q-S-R-S-L-(therapeutic agent) (SEQ ID NO: 551);  
Ac-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 362);  
Ac-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 363);  
20 Ac-S-G-R-S-S-S-L-(therapeutic agent) (SEQ ID NO: 364);  
Ac-S-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 365);  
Ac-S-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 366); isomer 1  
Ac-S-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 367); isomer 2  
Ac-S-G-R-S-G(hex)-(therapeutic agent) (SEQ ID NO: 368);  
25 Ac-S-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 369);  
Ac-S-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 370);  
Ac-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 371);  
Ac-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 372);  
Ac-S-S-R-S-nL-(therapeutic agent) (SEQ ID NO: 373);

-266-

- Ac-T-G-R-S-Abu-(therapeutic agent) (SEQ ID NO: 374);  
Ac-T-G-R-S-L-(therapeutic agent) (SEQ ID NO: 375);  
Ac-T-G-R-S-nV-(therapeutic agent) (SEQ ID NO: 376);  
Ac-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 377);  
5 Ac-T-G-R-S-G(hex)-(therapeutic agent) (SEQ ID NO: 378);  
Ac-T-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 379);  
Ac-T-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 380);  
Ac-T-G-R-T-Abu-(therapeutic agent) (SEQ ID NO: 381);  
Ac-T-G-R-hS-nL-(therapeutic agent) (SEQ ID NO: 382);  
10 Ac-T-G-R-Abu-nL-(therapeutic agent) (SEQ ID NO: 383);  
Ac-T-G-R-Abu-nV-(therapeutic agent) (SEQ ID NO: 384);  
Ac-T-G-F(Gn)-S-nL-(therapeutic agent) (SEQ ID NO: 385);  
Ac-T-G-F(Gn)-S-Cha-(therapeutic agent) (SEQ ID NO: 386);  
Ac-T-G-F(Gn)-Abu-nV-(therapeutic agent) (SEQ ID NO: 387);  
15 Ac-T-G-K(alloc)-S-nL-(therapeutic agent) (SEQ ID NO: 388);  
Ac-T-G-K-S-nL-(therapeutic agent) (SEQ ID NO: 389);  
Ac-T-G-hR-S-nL-(therapeutic agent) (SEQ ID NO: 390);  
Ac-(hS)G-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 391);  
MeOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 392);  
20 PhSO<sub>2</sub>-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 393);  
MeOEtCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 394);  
MeO(EtO)<sub>2</sub>Ac-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 395);  
4-oxo-Pentanoyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 396);  
3,4-MethyldioxyPhAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 397);  
25 2-PyridylAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 398);  
PhOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 399);  
L-3-PhLactyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 400);  
MeOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 401);  
PhAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 402);

-267-

- MeOEtOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 403);  
MeOEtOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 404);  
HOOCButa-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 405);  
Z-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 406);
- 5 EtOCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 407);  
 $\beta$ A-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 408);  
Pent-4-ynoyl-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 409);  
NapAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 410);  
iBoc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 411);
- 10 HOAc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 412);  
MeSucc-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 413);  
N,N-diMeGly-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 414);  
Succ-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 415);  
HCO-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 416);
- 15 Ac-T-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 417);  
Ac-T-A-F(Gn)-S-nL-(therapeutic agent) (SEQ ID NO: 418);  
Ac-T-A-R-Abu-nV-(therapeutic agent) (SEQ ID NO: 419);  
Ac-T-A-R-S-Abu-(therapeutic agent) (SEQ ID NO: 420);  
Ac-T-A-R-T-Abu-(therapeutic agent) (SEQ ID NO: 421);
- 20 Ac-T-S(O-Me)-R-S-nL-(therapeutic agent) (SEQ ID NO: 422);  
Ac-T-hS-R-S-nL-(therapeutic agent) (SEQ ID NO: 423);  
Ac-T-(1-Me)H-R-S-nL-(therapeutic agent) (SEQ ID NO: 424);  
Ac-T-(3-Me)H-R-S-nL-(therapeutic agent) (SEQ ID NO: 425);  
Ac-T-H-R-S-nL-(therapeutic agent) (SEQ ID NO: 426);
- 25 Ac-T-Sar-R-S-nL-(therapeutic agent) (SEQ ID NO: 427);  
Ac-T-nV-R-S-nL-(therapeutic agent) (SEQ ID NO: 428);  
Ac-T-nL-R-S-nL-(therapeutic agent) (SEQ ID NO: 429);  
Ac-T-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 430);  
Ac-T-Abu-R-S-nL-(therapeutic agent) (SEQ ID NO: 431);



-268-

- Ac-4,4diMeThr-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 432);  
Ac-hS-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 433);  
Ac-hS-G-R-hS-Cha-(therapeutic agent) (SEQ ID NO: 434);  
Ac-hS-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 435);  
5 Ac-hS-G-R-T-Cha-(therapeutic agent) (SEQ ID NO: 436);  
Ac-hS-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 437);  
Ac-N-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 438);  
Ac-Y-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 439);  
Ac-Y-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 440);  
10 Ac-Q-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 441);  
Ac-Q-G-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 442);  
Ac-L-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 573);  
Ac-L-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 342);  
Ac-L-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 343);  
15 Ac-L-R-G-S-G-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 344);  
Ac-L-R-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 345);  
Ac-L-R-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 574);  
Ac-L-R-G-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 346 );  
Ac-L-R-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 347);  
20 Ac-L-R-G-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 348);  
Ac-L-R-G-S-A-R-S-S-nV-(therapeutic agent) (SEQ ID NO: 349);  
Ac-L-R-G-S-A-R-S-S-nL-(therapeutic agent) (SEQ ID NO: 350);  
Ac-V-I-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 351);  
Ac-V-I-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 352);  
25 Ac-V-I-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 353);  
Ac-V-I-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 354);  
Ac-V-I-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 355);  
Ac-V-I-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 356);  
Ac-V-I-V-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 357);

-269-

- Ac-V-I-V-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 358);  
Ac-V-I-V-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 359);  
Ac-R-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 360);  
Ac-R-R-nV-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 361);
- 5 Ac-R-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 309);  
Ac-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 310);  
Ac-R-G-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 311);  
Ac-R-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 312);  
Ac-R-G-S-G-R--S-nL-(therapeutic agent) (SEQ ID NO: 313);
- 10 Ac-R-G-S-G-R-A-nL-(therapeutic agent) (SEQ ID NO: 314);  
Ac-R-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 315);  
Ac-R-G-S-G-R-S-Cha-(therapeutic agent) (SEQ ID NO: 316);  
Ac-R-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 317);  
Ac-R-G-S-A-R-S-Cha-(therapeutic agent) (SEQ ID NO: 318);
- 15 Ac-R-G-S-A-R-S-S-(therapeutic agent) (SEQ ID NO: 319);  
Ac-R-G-S-A-R-S-nV-(therapeutic agent) (SEQ ID NO: 320);  
Ac-R-G-S-A-R-S-S-nV -(therapeutic agent) (SEQ ID NO: 321);  
Ac-R-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 322);  
Ac-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 323);
- 20 Ac-R-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 324);  
Ac-R-C(Me)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 325);  
Ac-R-L-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 326);  
Ac-R-V-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 327);  
Ac-R-V-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 328);
- 25 Ac-R-nL-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 329);  
Ac-R-G(tBu)-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 330);  
Ac-R-L-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 331);  
Ac-R-V-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 332);  
Ac-R-nL-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 333);

-270-

- Ac-I-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 334);  
 Ac-I-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 335);  
 Ac-I-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 336);  
 Ac-I-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 337);  
 5 Ac-I-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 338);  
 Ac-I-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 339);  
 Ac-I-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 340);  
 Ac-I-V-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 341);  
 Ac-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 585);  
 10 Ac-G-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 277);  
 Ac-G-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 278);  
 Ac-G-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 279);  
 Ac-G-S-G-R-L-(therapeutic agent) (SEQ ID NO: 280);  
 Ac-G-S-G-(4-guan)Phg-S-L-(therapeutic agent) (SEQ ID NO: 281);  
 15 Ac-G-S-G-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 282);  
 Ac-G-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 283);  
 Ac-G-S-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 284);  
 Ac-G-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 285);  
 Succ-bA-T-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 286);  
 20 Ac-G-T-G-R-S-hCha-(therapeutic agent) (SEQ ID NO: 287);  
 Ac-G-hS-G-R-S-nL-(therapeutic agent) (SEQ ID NO: 288);  
 Ac-G-dS-A-R-S-A-(therapeutic agent) (SEQ ID NO: 289);  
 Ac-G-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 290);  
 Ac-G-S-A-R-S-S-Cha-(therapeutic agent) (SEQ ID NO: 291);  
 25 Ac-G-S-A-R-S-S-L-(therapeutic agent) (SEQ ID NO: 292);  
 Ac-G-S-A-R-A-S-L-(therapeutic agent) (SEQ ID NO: 293);  
 Ac-V-S-G-R-S-L-(therapeutic agent) (SEQ ID NO: 294);  
 Ac-V-S-G-R-A-L-(therapeutic agent) (SEQ ID NO: 295);  
 Ac-V-S-G-R-A-S-L-(therapeutic agent) (SEQ ID NO: 296);

-271-

- Ac-V-S-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 297);  
Ac-V-S-A-R-M-A-(therapeutic agent) (SEQ ID NO: 298);  
Ac-V-S-A-R-nL-A-(therapeutic agent) (SEQ ID NO: 299);  
Ac-V-S-A-R-S-nL-(therapeutic agent) (SEQ ID NO: 300);  
5 Ac-V-S-A-R-S-L-(therapeutic agent) (SEQ ID NO: 301);  
Ac-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 302);  
Ac-(Me)C-P-G-R-V-V-(therapeutic agent) (SEQ ID NO: 303);  
Ac-C(Me)-P-G-R-A-L-(therapeutic agent) (SEQ ID NO: 304);  
Ac-C(Me)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 305);  
10 Ac-C(Me)-P-A-R-S-L-(therapeutic agent) (SEQ ID NO: 306);  
Ac-C(Me)-P-A-R-A-S-L-(therapeutic agent) (SEQ ID NO: 307);  
Ac-G(tBu)-P-G-R-S-L-(therapeutic agent) (SEQ ID NO: 308);  
Ac-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 552);  
Ac-Q-S-R-S-A-(therapeutic agent) (SEQ ID NO: 553);  
15 Ac-Q-S-R-S-G-(therapeutic agent) (SEQ ID NO: 554);  
Ac-R-S-R-A-A-(therapeutic agent) (SEQ ID NO: 555);  
Ac-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 556);  
Ac-R-Q-S-R-S-A-(therapeutic agent) (SEQ ID NO: 557);  
Ac-R-Q-S-R-S-A-A-(therapeutic agent) (SEQ ID NO: 558);  
20 Ac-R-G-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 559);  
Ac-S-G-R-A-A-(therapeutic agent) (SEQ ID NO: 560);  
Ac-S-G-R-S-A-(therapeutic agent) (SEQ ID NO: 561);  
Ac-S-G-R-S-S-A-(therapeutic agent) (SEQ ID NO: 562);  
Ac-S-G-R-A-S-A-(therapeutic agent) (SEQ ID NO: 563);  
25 Ac-S-G-R-S-G-(therapeutic agent) (SEQ ID NO: 564);  
Ac-S-G-R-S-S-G-(therapeutic agent) (SEQ ID NO: 565);  
Ac-S-G-R-S-G-A-(therapeutic agent) (SEQ ID NO: 566);  
Ac-S-G-R-S-G-G-(therapeutic agent) (SEQ ID NO: 567);  
Ac-G-T-G-R-S-G-G-(therapeutic agent) (SEQ ID NO: 568);

-272-

- Ac-G-S-G-R-S-G-G-(therapeutic agent) (SEQ ID NO: 243)  
 Ac-L-R-R-Q-S-R-A-A-(therapeutic agent) (SEQ ID NO: 597);  
 MeSO<sub>2</sub>-dA(Chx)-Abu-R-S-L-(therapeutic agent) (SEQ ID NO: 598);  
 Ac-R-A-R-S-L-(therapeutic agent) (SEQ ID NO: 599);
- 5 Ac-dA(Chx)-Abu-R-S-L-(therapeutic agent) (SEQ ID NO: 600);  
 Ac-dA(Chx)-Abu-R-S-S-L-(therapeutic agent) (SEQ ID NO: 601);  
 Ac-Q-G-R-S-S-L-(therapeutic agent) (SEQ ID NO: 602);  
 MeOCO-dhF-P(OH)-R-S-S-L-(therapeutic agent) (SEQ ID NO: 603);  
 MeOCO-Quat4-G-R-S-L-(therapeutic agent) (SEQ ID NO: 604);
- 10 Ac-dCha-P(OH)-R-S-S-L-(therapeutic agent) (SEQ ID NO: 605);  
 Ac-dCha-Abu-R-S-S-A-(therapeutic agent) (SEQ ID NO: 606);  
 MeOCO-Quat2-G-R-S-L-(therapeutic agent) (SEQ ID NO: 607);  
 MeOCO-Quat3-G-R-S-L-(therapeutic agent) (SEQ ID NO: 608); and  
 MeOCO-Quat-G-R-S-L-(therapeutic agent) (SEQ ID NO: 609).
- 15 105. The conjugate of any of claims 35-56, wherein P<sub>4</sub> is selected from Pro, Arg, Ser, Ala, Lys, Gly, nLeu, Leu, Tyr, Glu, Phe, Val, N,N-dimethylGly,  $\beta$ -Ala, Cys(Me), Gln, t-butylGly and nVal.
106. The conjugate of claim 1 or 66, that comprises a peptidic substrate of the formula P<sub>6</sub>-P<sub>5</sub>-P<sub>4</sub>-P<sub>3</sub>-P<sub>2</sub>-P<sub>1</sub>-P<sub>1'</sub>-P<sub>2'</sub>-P<sub>3'</sub>-P<sub>4'</sub>, wherein
- 20 each of P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, P<sub>4</sub>, P<sub>5</sub>, P<sub>6</sub>, P<sub>1'</sub> and P<sub>2'</sub> are selected from residues set forth in Figures 1 and 2, and P<sub>6</sub>, P<sub>5</sub>, P<sub>4</sub>, P<sub>2'</sub>, P<sub>3'</sub> and P<sub>4'</sub> are optional.
107. The conjugate of claim 67, wherein:
- 25 P<sub>6</sub> is optional and is selected from L, V, R;  
 P<sub>5</sub> is optional and is selected from R, I, L;  
 P<sub>4</sub> is optional and is selected from G, C, V;  
 P<sub>3</sub> is selected from S, dS, P, A or G;  
 P<sub>2</sub> is selected from A or G;  
 P<sub>1</sub> is R;

-273-

P1' is S, V, M or nL;

P2' is optional and is selected S, L, A or V;

P3' is optional and is L; and

P4' is optional and is L.

5           108. A conjugate, comprising a therapeutic agent and a nucleic acid substrate linked thereto via a peptidic linker, wherein the peptidic linker is proteolytically cleaved by a cell surface protease or a soluble, released or shed form thereof, to liberate the therapeutic agent, wherein the conjugate is not substantially cleaved by plasmin or prostate specific  
10   antigen (PSA).

          109. The conjugate of claim 108, wherein the nucleic acid is DNA.

          110. The conjugate of claim 108, wherein the nucleic acid is RNA.

15           111. The conjugate of claim 108, wherein the nucleic acid is double-stranded RNA.

          112. The conjugate of claim 67, wherein:

          P6 is optional and is selected from L, V, R;

          P5 is optional and is selected from R, I, L;

20           P4 is optional and is selected from G, C, V;

          P3 is selected from S, dS, P, A or G;

          P2 is selected from A or G;

          P1 is R;

          P1' is T, Abu, hS, nV or A;

25           P2' is optional and is selected S, L, A or V;

          P3' is optional and is L, nL, nV, G(hex), G(allyl), CHA, hCHA, or Abu; and

          P4' is optional and is L, nL, nV, G(hex), G(allyl), CHA, hCHA, or Abu.

-274-

113. The conjugate of claim 67, wherein:

P6 is optional and is selected from L, V, R;

P5 is optional and is selected from R, I, L;

P4 is optional and is selected from G, C, V;

5 P3 is selected from S, dS, P, A or G;

P2 is selected from A or G;

P1 is R;

P1' is S, G or A;

P2' is optional and is selected G or A;

10 P3' is optional and is L, nL, nV, G(hex), G(allyl), CHA, hCHA,  
or Abu; and

P4' is optional and is L, nL, nV, G(hex), G(allyl), CHA, hCHA,  
or Abu.

114. The conjugate of any of claims 1-113, wherein the  
15 therapeutic agent is taxol.

115. The conjugate of any of claims 1-113, wherein the  
therapeutic agent is doxorubicin.

116. A method of treatment of a disease, comprising  
administering a conjugate of any of claims 1-113 to a subject, wherein  
20 the disease is a cell-surface protease-associated disease.

117. The method of claim 116, wherein the disease is selected  
from the group consisting of autoimmune diseases, inflammatory  
diseases, infectious diseases and endocrine diseases.

118. The method of claim 116, wherein the disease is a  
25 proliferative disease.

119. A method of treatment of a cell-surface protease-associated  
disease, comprising administering a conjugate, comprising a therapeutic  
agent and a peptidic substrate linked thereto optionally via a linker,  
wherein the peptidic substrate is proteolytically cleaved by a cell surface

-275-

protease or a soluble, released or shed form thereof to liberate the therapeutic agent, to a subject exhibiting symptoms of a cell-surface protease-associated disorder.

120. The method of claim 119, wherein the disease is selected  
5 from the group consisting of autoimmune diseases, inflammatory diseases, infectious diseases and endocrine diseases.

121. The method of claim 119, wherein the disease is a proliferative disease.

122. The method of any of claims 114-119, wherein the subject  
10 is a mammal.

123. The method of claim 120, wherein the mammal is a human.

124. The method of claim 118 or 121, wherein the disease is cancer.

125. The method of claim 118 or 121, wherein the disease is  
15 selected from ocular disorders, cardiovascular disorders, chronic inflammatory diseases, wounds, circulatory disorders, dermatological disorders and cancer.

126. The method of claim 118 or 121, wherein the disease is selected from rheumatoid arthritis, psoriasis, diabetic retinopathies,  
20 recurrence of pterygii, scarring from excimer laser surgery, scarring from glaucoma filtering surgery, macular degeneration anterior eye, crest syndromes, solid neoplasms and vascular tumors.

127. The method of claim 118 or 121, wherein the disease is selected from lung cancer, colon cancer, pancreatic cancer, esophageal  
25 cancer, breast cancer, ovarian cancer, prostate cancer, melanoma and Kaposi's sarcoma.

128. The method of any of claims 116-127, wherein the therapeutic agent is taxol.



-276-

129. The method of any of claims 116-127, wherein the therapeutic agent is doxorubicin.

130. A pharmaceutical composition, comprising the conjugate of any of claims 1-113 or a pharmaceutically acceptable derivative thereof,  
5 in a pharmaceutically acceptable carrier.

131. The pharmaceutical composition of claim 130 that is formulated for single dosage administration.

132. An article of manufacture, comprising packaging material, the conjugate of any of claims 1-113, or a pharmaceutically acceptable  
10 derivative thereof, contained within packaging material, which is used for treatment, prevention or amelioration of one or more symptoms associated with cell-surface protease-associated diseases or disorders, and a label that indicates that the conjugate or pharmaceutically acceptable derivative thereof is used for treatment, prevention or  
15 amelioration of one or more symptoms associated with cell-surface protease-associated diseases or disorders.

133. The conjugate of any of claims 1-113 when used for the treatment of a cell-surface protease-associated disease.

134. The conjugate of claim 133, wherein the disease is a  
20 proliferative disease.

135. The conjugate of claim 134, wherein the proliferative disease is cancer.

136. The conjugate of claim 134, wherein the proliferative disease is selected from ocular diseases, cardiovascular diseases, chronic  
25 inflammatory diseases, wounds, circulatory diseases, dermatological diseases and cancer.

137. The conjugate of claim 134, wherein the proliferative disease is selected from rheumatoid arthritis, psoriasis, diabetic retinopathies, recurrence of pterygii, scarring from excimer laser surgery, scarring from

-277-

glaucoma filtering surgery, macular degeneration anterior eye, crest syndromes, solid neoplasms and vascular tumors.

138. The conjugate of claim 134, wherein the proliferative disease is selected from lung cancer, colon cancer, pancreatic cancer, esophageal  
5 cancer, breast cancer, ovarian cancer, prostate cancer, melanoma and Kaposi's sarcoma.

139. Use of the conjugate of any of claims 1-113 for the preparation of a medicament for use in the treatment of a cell-surface protease-associated disease.

10 140. The use of claim 139, wherein the disease is a proliferative disease.

141. The use of claim 140, wherein the proliferative disease is cancer.

142. The use of claim 140, wherein the proliferative disease is  
15 selected from ocular diseases, cardiovascular diseases, chronic inflammatory diseases, wounds, circulatory diseases, dermatological diseases and cancer.

143. The use of claim 140, wherein the proliferative disease is selected from rheumatoid arthritis, psoriasis, diabetic retinopathies,  
20 recurrence of pterygii, scarring from excimer laser surgery, scarring from glaucoma filtering surgery, macular degeneration anterior eye, crest syndromes, solid neoplasms and vascular tumors.

144. The use of claim 140, wherein the proliferative disease is selected from lung cancer, colon cancer, pancreatic cancer, esophageal  
25 cancer, breast cancer, ovarian cancer, prostate cancer, melanoma and Kaposi's sarcoma.

145. A method of preparing a conjugate of any of claims 1-113, comprising:

a) synthesizing the peptidic substrate;

-278-

- b) optionally capping the peptidic substrate on either the N-terminus or the C-terminus;
- c) optionally linking the non-capped terminus of the peptidic substrate to a linker;
- 5 d) coupling the peptidic substrate to a therapeutic agent, optionally via the linker, to form a conjugate; and
- e) optionally, deprotecting the conjugate, if protected.

146. The method of claim 145, wherein, prior to step a), the method comprises a step of identifying a peptidic substrate for the  
10 protease.

147. A method, comprising:
- a) selecting a disease;
  - b) identifying a cell involved in the disease process or a cell in the vicinity of the cell involved in the disease process; and
  - 15 c) identifying a cell surface protease on the cell, thereby identifying proteases to target conjugates for treatment of diseases.

148. The method of claim 147, further comprising preparing a conjugate that targets the protease.

CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-Y-G-R-S-S-L-Dox	B	481
Ac-M(O2)-S-R-S-L-Dox	C	482
Ac-R-R-Q-S-R-A-A-Dox	A	105
Ac-R-R-Q-S-R-I-Dox; isomer 1	D	610
Ac-R-R-Q-S-R-S-S-L-Dox	A	543
Ac-R-R-Q-S-R-S-L-Dox	A	544
Ac-R-G-S-G-R-S-L-Dox	B	545
Ac-R-G-S-G-R-S-nL-Dox	A	546
Ac-R-G-S-G-R-A-nL-Dox	A	547
Ac-R-G-S-G-R-S-S-L-Dox	B	548
Ac-I-V-S-G-R-A-S-L-Dox	C	549
Ac-R-R-Q-S-R-A-Dox	NT	108
Ac-R-R-Q-S-R-I-Dox; isomer 2	NT	111
Ac-L-R-R-Q-S-R-A-A-Dox	A	106
Ac-L-R-R-Q-S-R-G-G-Dox	B	109
Ac-L-R-R-Q-S-R-A-Dox	C	110
Ac-L-R-R-Q-S-R-A-I-Dox; isomer 1	A	112
Ac-L-R-R-Q-S-R-A-I-Dox; isomer 2	C	611
Ac-L-R-R-Q-S-R-S-S-I-Dox	A	550
Ac-L-R-R-Q-S-R-S-L-Dox	A	551

FIG. 1A

CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-Q-S-R-S-S-nV-Dox	B	457
Ac-Q-S-R-S-S-Cha-Dox	B	458
Ac-Q-S-R-S-S-L-Dox	B	459
Ac-Q-T-R-S-S-L-Dox	C	460
Ac-Q-Aib-R-S-S-Cha-Dox	C	461
Ac-Q-Aib -R-S-S-L-Dox	D	462
Ac-Q-Abu-R-S-S-Cha-Dox	B	463
Ac-Q-Abu-R-S-S-L-Dox	B	464
Ac-Q-Cha-R-S-S-Cha-Dox	D	465
Ac-Q-F-R-S-L-Dox	C	466
Ac-Q-F-R-S-S-L-Dox	B	467
Ac-Q-Y-R-S-S-L-Dox	C	468
Ac-R-G-R-S-L-Dox	A	469
Ac-R-G-R-S-S-L-Dox	A	470
Ac-R-G-R-S-S-Cha-Dox	A	471
Ac-R-G-R-S-Cha-Dox	A	472
Ac-R-A-R-S-L-Dox	B	473
Ac-R-A-R-S-S-L-Dox	A	474
Ac-R-S-R-S-L-Dox	B	475
Ac-R-S-R-S-S-L-Dox	B	476
Ac-R-S-R-S-Cha-Dox	B	477
Ac-R-S-R-S-S-Cha-Dox	B	478
Ac-R-F-R-S-L-Dox	B	479
Ac-R-F-R-S-Cha-Dox	B	480

FIG. 1B

CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-R-Q-F-R-A-nV-Dox	B	540
Ac-R-Q-F-R-A-Cha-Dox	D	541
Ac-Q-S-R-S-S-nL-Dox	B	542
MeOCO-Quat2-G-R-S-L-NH2	B	483
MeOCO-Quat3-G-R-S-L-NH2	B	484
MeOCO-Quat-G-R-S-L-NH2	C	485
MeOCO-Quat4-G-R-S-L-NH2	B	486
MeOCO-Quat5-G-R-S-L-NH2	C	487
MeOCO-Quat2-G-R-S-S-L-NH2	B	488
MeOCO-Quat4-G-R-S-L-Dox	B	489
MeOCO-Quat2-G-R-S-L-Dox	B	490
Ac-Q-G-R-S-L-Dox	C	445
Ac-Q-G-R-S-S-L-Dox	B	446
Ac-Q-G-R-A-S-L-Dox	B	447
Ac-N-G-R-S-S-L-Dox	C	448
Ac-Q-G-R-S-S-nL-Dox	B	449
Ac-Q-G-R-S-S-nV-Dox	B	450
Ac-Q-G-R-S-S-Cha-Dox	B	451
Ac-Q-G-R-S-S-allylG-Dox; isomer 1	B	452
Ac-Q-G-R-S-S-allylG-Dox; isomer 2	B	453
Ac-Q-A-R-S-L-Dox	C	454
Ac-Q-A-R-S-S-L-Dox	B	455
Ac-Q-S-R-S-L-Dox	C	456

FIG. 1C

CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-R-Q-S-R-A-A-Dox	A	515
Ac-R-Q-S-R-A-Dox	D	516
Ac-R-Q-S-R-A-nL-Dox	A	517
Ac-R-Q-S-R-A-L-Dox	A	519
Ac-R-Q-S-R-A-nV-Dox	A	520
Ac-R-Q-S-R-A-Cha-Dox	B	521
Ac-R-Q-S-R-S-S-L-Dox	A	522
Ac-R-Q-S-R-S-L-Dox	A	523
Ac-R-Q-S-R-S-dnL-Dox	A	524
Ac-R-Q-S-R-S-dnL-Dox	C	525
Ac-R-Q-S-R-S-nV-Dox	A	526
Ac-R-Q-S-R-S-allylG-Dox	A	527
Ac-R-Q-S-R-S-Cha-Dox	B	528
Ac-R-Q-S-R-T-nL-Dox	A	529
Ac-R-Q-T-R-S-S-L-Dox	B	530
Ac-R-Q-T-R-S-L-Dox	B	531
Ac-R-N-S-R-S-nL-Dox	B	532
Ac-R-Q-F-R-S-L-Dox	B	533
Ac-R-Q-F-R-S-nL-Dox	A	534
Ac-R-Q-F-R-S-nV-Dox	A	535
Ac-R-Q-F-R-S-nL-Dox	D	536
Ac-R-Q-F-R-S-Cha-Dox	D	537
Ac-R-Q-F-R-A-L-Dox	C	538
Ac-R-Q-F-R-A-nL-Dox	C	539

FIG. 1D

CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-R-Q-G-R-S-L-Dox	A	491
Ac-R-Q-G-R-S-S-L-Dox	A	492
Ac-R-Q-G-R-S-nL-Dox	A	493
Ac-R-Q-G-R-S-nV-Dox	A	494
Ac-R-Q-G-R-S-F-Dox	A	495
Ac-R-Q-G-R-A-L-Dox	A	496
Ac-R-Q-G-R-A-dL-Dox	D	497
Ac-R-Q-G-R-A-dnL-Dox	A	498
Ac-R-Q-G-R-A-nL-Dox	B	499
Ac-R-Q-G-R-A-nV-Dox	A	500
Ac-R-Q-G-R-A-Cha-Dox	A	501
Ac-R-Q-G-R-A-F-Dox	D	502
Ac-R-N-G-R-S-L-Dox	B	503
Ac-R-N-G-R-A-nL-Dox	A	504
Ac-R-Q-A-R-S-L-Dox	B	505
Ac-R-Q-A-R-S-nL-Dox	A	506
Ac-R-Q-A-R-S-nV-Dox	A	507
Ac-R-Q-A-A-S-Cha-Dox	B	508
Ac-R-Q-A-R-S-S-Cha-Dox	A	509
Ac-R-Q-A-R-T-nL-Dox	A	510
Ac-R-Q-A-R-A-L-Dox	D	511
Ac-R-Q-A-R-A-nL-Dox	A	512
Ac-R-Q-A-R-A-nV-Dox	B	513
Ac-R-Q-A-R-A-Cha-Dox	B	514

FIG. 1E



CONJUGATE	uPA CT50	SEQ ID NO
Ac-S-G-R-S-L-Dox	C	362
Ac-S-G-R-S-S-L-Dox	C	363
Ac-S-G-R-S-S-S-L-Dox	D	364
Ac-S-G-R-S-nL-Dox	B	365
Ac-S-G-R-S-nV-Dox; isomer 1	B	366
Ac-S-G-R-S-nV-Dox; isomer 2	D	367
Ac-S-G-R-S-G(hex)-Dox	A	368
Ac-S-G-R-S-Cha-Dox	B	369
Ac-S-G-R-S-hCha-Dox	A	370
Ac-S-A-R-S-L-Dox	D	371
Ac-S-A-R-S-S-L-Dox	D	372
Ac-S-S-R-S-nL-Dox	C	373
Ac-T-G-R-S-Abu-Dox	A	374
Ac-T-G-R-S-L-Dox	B	375
Ac-T-G-R-S-nV-Dox	A	376
Ac-T-G-R-S-nL-Dox	A	377
Ac-T-G-R-S-G(hex)-Dox	A	378
Ac-T-G-R-S-Cha-Dox	B	379
Ac-T-G-R-S-hCha-Dox	A	380
Ac-T-G-R-T-Abu-Dox	B	381
Ac-T-G-R-hS-nL-Dox	B	382
Ac-T-G-R-Abu-nL-Dox	A	383
Ac-T-G-R-Abu-nV-Dox	B	384
Ac-T-G-F(Gn)-S-nL-Dox	A	385

FIG. 2A

CONJUGATE	uPA CT50	SEQ ID NO
Ac-T-G-F(Gn)-S-Cha-Dox	A	386
Ac-T-G-F(Gn)-Abu-nV-Dox	NT	387
Ac-T-G-K(alloc)-S-nL-Dox	D	388
Ac-T-G-K-S-nL-Dox	B	389
Ac-T-G-hR-S-nL-Dox	D	390
Ac-(hS)G-G-R-S-nL-Dox	D	391
MeOCO-T-G-R-S-nL-Dox	A	392
PhSO <sub>2</sub> -T-G-R-S-nL-Dox	B	393
MeOEtCO-T-G-R-S-nL-Dox	A	394
MeO(EtO) <sub>2</sub> Ac-T-G-R-S-nL-Dox	A	395
4-oxo-Pentanoyl-T-G-R-S-nL-Dox	A	396
3,4-MethyldioxyPhAc-T-G-R-S-nL-Dox	A	397
2-PyridylAc-T-G-R-S-nL-Dox	A	398
PhOAc-T-G-R-S-nL-Dox	A	399
L-3-PhLactyl-T-G-R-S-nL-Dox	A	400
MeOAc-T-G-R-S-nL-Dox	A	401
PhAc-T-G-R-S-nL-Dox	A	402
MeOEtOCO-T-G-R-S-nL-Dox	A	403
MeOEtOAc-T-G-R-S-nL-Dox	A	404
HOOCButa-T-G-R-S-nL-Dox	A	405
Z-T-G-R-S-nL-Dox	A	406
EtOCO-T-G-R-S-nL-Dox	A	407
$\beta$ A-T-G-R-S-nL-Dox	A	408
Pent-4-ynoyl-T-G-R-S-nL-Dox	A	409

FIG. 2B

CONJUGATE	uPA CT50	SEQ ID NO
NapAc-T-G-R-S-nL-Dox	B	410
iBoc-T-G-R-S-nL-Dox	A	411
HOAc-T-G-R-S-nL-Dox	A	412
MeSucc-T-G-R-S-nL-Dox	A	413
N,N-diMeGly-T-G-R-S-nL-Dox	A	414
Succ-T-G-R-S-nL-Dox	B	415
HCO-T-G-R-S-nL-Dox	A	416
Ac-T-A-R-S-nL-Dox	A	417
Ac-T-A-F(Gn)-S-nL-Dox	A	418
Ac-T-A-R-Abu-nV-Dox	NT	419
Ac-T-A-R-S-Abu-Dox	B	420
Ac-T-A-R-T-Abu-Dox	B	421
Ac-T-S(O-Me)-R-S-nL-Dox	B	422
Ac-T-hS-R-S-nL-Dox	B	423
Ac-T-(1-Me)H-R-S-nL-Dox	NT	424
Ac-T-(3-Me)H-R-S-nL-Dox	NT	425
Ac-T-H-R-S-nL-Dox	C	426
Ac-T-Sar-R-S-nL-Dox	D	427
Ac-T-nV-R-S-nL-Dox	D	428
Ac-T-nL-R-S-nL-Dox	B	429
Ac-T-A-R-S-Cha-Dox	B	430
Ac-T-Abu-R-S-nL-Dox	B	431
Ac-4,4diMeThr-G-R-S-nL-Dox	B	432
Ac-hS-G-R-S-nL-Dox	D	433

FIG. 2C

CONJUGATE	uPA CT50	SEQ ID NO
Ac-hS-G-R-hS-Cha-Dox	D	434
Ac-hS-G-R-S-Cha-Dox	D	435
Ac-hS-G-R-T-Cha-Dox	D	436
Ac-hS-A-R-S-Cha-Dox	D	437
Ac-N-G-R-S-nL-Dox	D	438
Ac-Y-G-R-S-S-L-Dox	D	439
Ac-Y-G-R-S-Cha-Dox	D	440
Ac-Q-G-R-S-S-nL-Dox	D	441
Ac-Q-G-R-S-S-nV-Dox	D	442
Ac-L-R-G-S-G-R-S-A-Dox	B	573
Ac-L-R-G-S-G-R-S-L-Dox	C	342
Ac-L-R-G-S-G-R-S-dL-Dox	D	343
Ac-L-R-G-S-G-R-S-S-nL-Dox	D	344
Ac-L-R-G-S-G-R-S-S-Cha-Dox	C	345
Ac-L-R-G-dS-A-R-S-A-Dox	C	574
Ac-L-R-G-S-A-R-S-S-L-Dox	D	346
Ac-L-R-G-S-A-R-S-L-Dox	C	347
Ac-L-R-G-S-A-R-S-S-Cha-Dox	C	348
Ac-L-R-G-S-A-R-S-S-nV-Dox	D	349
Ac-L-R-G-S-A-R-S-S-nL-Dox	D	350
Ac-V-I-V-S-G-R-A-L-Dox	D	351
Ac-V-I-V-S-A-R-S-L-Dox	D	352
Ac-V-I-V-S-G-R-S-S-L-Dox	C	353
Ac-V-I-V-S-A-R-M-A-Dox	C	354

FIG. 2D

CONJUGATE	uPA CT50	SEQ ID NO
Ac-V-I-V-S-A-R-nL-A-Dox	D	355
Ac-V-I-V-S-A-R-S-nL-Dox	D	356
Ac-V-I-V-S-A-R-S-Cha-Dox	D	357
Ac-V-I-V-S-A-R-S-dCha-Dox	D	358
Ac-V-I-V-S-A-R-S-S-Cha-Dox	D	359
Ac-R-R-(Me)C-P-G-R-V-V-Dox	D	360
Ac-R-R-nV-P-A-R-S-L-Dox	D	361
Ac-R-G-dS-A-R-S-A-Dox	C	309
Ac-R-G-S-G-R-S-A-Dox	A	310
Ac-R-G-S-G-R-A-L-Dox	D	311
Ac-R-G-S-G-R-S-L-Dox	B	312
Ac-R-G-S-G-R--S-nL-Dox	A	313
Ac-R-G-S-G-R-A-nL-Dox	B	314
Ac-R-G-S-G-R-S-S-L-Dox	C	315
Ac-R-G-S-G-R-S-Cha-Dox	C	316
Ac-R-G-S-G-R-S-S-Cha-Dox	C	317
Ac-R-G-S-A-R-S-Cha-Dox	B	318
Ac-R-G-S-A-R-S-S-Cha-Dox	B	319
Ac-R-G-S-A-R-S-nV-Dox	B	320
Ac-R-G-S-A-R-S-S-nV-Dox	C	321
Ac-R-G-S-A-R-S-L-Dox	D	322
Ac-R-(Me)C-P-G-R-V-V-Dox	D	323
Ac-R-(Me)C-P-G-R-V-V-Dox	D	324
Ac-R-C(Me)-P-G-R-S-L-Dox	D	325

FIG. 2E

CONJUGATE	uPA CT50	SEQ ID NO
Ac-R-L-P-G-R-S-L-Dox	D	326
Ac-R-V-P-G-R-S-L-Dox	D	327
Ac-R-V-P-G-R-S-dL-Dox	D	328
Ac-R-nL-P-G-R-S-L-Dox	D	329
Ac-R-G(tBu)-P-A-R-S-L-Dox	D	330
Ac-R-L-P-A-R-S-L-Dox	D	331
Ac-R-V-P-A-R-S-L-Dox	D	332
Ac-R-nL-P-A-R-S-L-Dox	D	333
Ac-I-V-S-G-R-A-L-Dox	D	334
Ac-I-V-S-G-R-S-S-L-Dox	D	335
Ac-I-V-S-G-R-A-S-L-Dox	D	336
Ac-I-V-S-A-R-M-A-Dox	B	337
Ac-I-V-S-A-R-nL-A-Dox	B	338
Ac-I-V-S-A-R-S-L-Dox	C	339
Ac-I-V-S-A-R-S-nL-Dox	B	340
Ac-I-V-S-A-R-S-S-L-Dox	C	341
Ac-G-S-G-R-S-A-Dox	B	585
Ac-G-S-G-R-S-L-Dox	C	277
Ac-G-S-G-R-A-L-Dox	D	278
Ac-G-S-G-R-S-S-L-Dox	D	279
Ac-G-S-G-R-L-Dox	D	280
Ac-G-S-G-(4-guan)Phg-S-L-NH <sub>2</sub>	D	281
Ac-G-S-G-R-S-S-Cha-Dox	D	282
Ac-G-S-G-R-A-S-L-Dox	D	283

FIG. 2F

CONJUGATE	uPA CT50	SEQ ID NO
Ac-G-S-G-R-S-nL-Dox	A	284
Ac-G-T-G-R-S-nL-Dox	A	285
Succ- $\beta$ A-T-G-R-S-nL-Dox	A	286
Ac-G-T-G-R-S-hCha-Dox	A	287
Ac-G-hS-G-R-S-nL-Dox	D	288
Ac-G-dS-A-R-S-A-Dox	C	289
Ac-G-S-A-R-S-L-Dox	D	290
Ac-G-S-A-R-S-S-Cha-Dox	C	291
Ac-G-S-A-R-S-S-L-Dox	D	292
Ac-G-S-A-R-A-S-L-Dox	D	293
Ac-V-S-G-R-S-L-Dox	D	294
Ac-V-S-G-R-A-L-Dox	D	295
Ac-V-S-G-R-A-S-L-Dox	D	296
Ac-V-S-G-R-S-S-L-Dox	D	297
Ac-V-S-A-R-M-A-Dox	B	298
Ac-V-S-A-R-nL-A-Dox	B	299
Ac-V-S-A-R-S-nL-Dox	B	300
Ac-V-S-A-R-S-L-Dox	D	301
Ac-(Me)C-P-G-R-V-dV-Dox	D	302
Ac-(Me)C-P-G-R-V-V-Dox	D	303
Ac-C(Me)-P-G-R-A-L-Dox	D	304
Ac-C(Me)-P-G-R-S-L-Dox	D	305
Ac-C(Me)-P-A-R-S-L-Dox	D	306
Ac-C(Me)-P-A-R-A-S-L-Dox	D	307

FIG. 2G

CONJUGATE	uPA CT50	SEQ ID NO
Ac-G(tBu)-P-G-R-S-L-Dox	D	308

FIG. 2H



CONJUGATE	MTSP1 CT50	SEQ ID NO
Ac-Q-S-R-A-A-Tax	B	552
Ac-Q-S-R-S-A-Tax	B	553
Ac-Q-S-R-S-G-Tax	C	554
Ac-R-S-R-A-A-Tax	B	555
Ac-R-Q-S-R-A-A-Tax	A	556
Ac-R-Q-S-R-S-A-Tax	A	557
Ac-R-Q-S-R-S-A-A-Tax	A	558
Ac-R-Q-S-R-A-A-Tax	A	559

FIG. 3

CONJUGATE	uPA CT50	SEQ ID NO
Ac-R-G-S-G-R-S-A-Tax	D	559
Ac-S-G-R-A-A-Tax	D	560
Ac-S-G-R-S-A-Tax	D	561
Ac-S-G-R-S-S-A-Tax	D	562
Ac-S-G-R-A-S-A-Tax	D	563
Ac-S-G-R-S-G-Tax	D	564
Ac-S-G-R-S-S-G-Tax	D	565
Ac-S-G-R-S-G-A-Tax	D	566
Ac-S-G-R-S-G-G-Tax	D	567
Ac-G-T-G-R-S-G-G-Tax	C	568
Ac-G-S-G-R-S-G-G-Tax	C	518

FIG. 4

CONJUGATE	ET1 CT50	SEQ ID NO
Ac-L-R-R-Q-S-R-A-A-Dox	B	597
MeSO <sub>2</sub> -dA(Chx)-Abu-R-S-L-Dox	D	598
Ac-R-A-R-S-L-Dox	B	599
Ac-dA(Chx)-Abu-R-S-L-Dox	C	600
Ac-dA(Chx)-Abu-R-S-S-L-Dox	B	601
Ac-Q-G-R-S-S-L-Dox	A	602
MeOCO-dhF-P(OH)-R-S-S-L-Dox	B	603
MeOCO-Quat4-G-R-S-L-Dox	D	604
Ac-dCha-P(OH)-R-S-S-L-Dox	B	605
Ac-dCha-Abu-R-S-S-A-Tax	B	606
MeOCO-Quat2-G-R-S-L-NH <sub>2</sub>	B	607
MeOCO-Quat3-G-R-S-L-NH <sub>2</sub>	B	608
MeOCO-Quat-G-R-S-L-NH <sub>2</sub>	C	609

FIG. 5

-1-

## SEQUENCE LISTING

<110> Edwin L. Madison  
Joseph Edward Semple  
George P. Vlasuk  
Scott Jeffrey Kemp  
Mallareddy Komandla  
Daniel Vanna Siev

<120> Conjugates Activated By Cell Surface Proteases and Therapeutic Uses Thereof

<130> 24745-1611PC

<140> Not Yet Assigned

<141> Herewith

<160> 611

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3147

<212> DNA

<213> Homo Sapien

<220>

<223> Nucleotide sequence encoding MTSP1

<221> CDS

<222> (23) ... (2589)

<300>

<301> O'Brien, T.J. and Tanimoto, H.

<308> GenBank AR081724

<310> US Pat 5972616

<311> 1998-02-20

<312> 1999-10-26

<400> 1

tcaagagcgg cctcggggta cc atg ggg agc gat cgg gcc cgc aag ggc gga	52
Met Gly Ser Asp Arg Ala Arg Lys Gly Gly	
1 5 10	
ggg ggc ccg aag gac ttc ggc gcg gga ctc aag tac aac tcc cgg cac	100
Gly Gly Pro Lys Asp Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His	
15 20 25	
gag aaa gtg aat ggc ttg gag gaa ggc gtg gag ttc ctg cca gtc aac	148
Glu Lys Val Asn Gly Leu Glu Glu Gly Val Glu Phe Leu Pro Val Asn	
30 35 40	
aac gtc aag aag gtg gaa aag cat ggc ccg ggg cgc tgg gtg gtg ctg	196
Asn Val Lys Lys Val Glu Lys His Gly Pro Gly Arg Trp Val Val Leu	
45 50 55	
gca gcc gtg ctg atc ggc ctc ctc ttg gtc ttg ctg ggg atc ggc ttc	244
Ala Ala Val Leu Ile Gly Leu Leu Leu Val Leu Leu Gly Ile Gly Phe	
60 65 70	
ctg gtg tgg cat ttg cag tac cgg gac gtg cgt gtc cag aag gtc ttc	292
Leu Val Trp His Leu Gln Tyr Arg Asp Val Arg Val Gln Lys Val Phe	
75 80 85 90	

-2-

aat ggc tac atg agg atc aca aat gag aat ttt gtg gat gcc tac gag Asn Gly Tyr Met Arg Ile Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu 95 100 105	340
aac tcc aac tcc act gag ttt gta agc ctg gcc agc aag gtg aag gac Asn Ser Asn Ser Thr Glu Phe Val Ser Leu Ala Ser Lys Val Lys Asp 110 115 120	388
gcg ctg aag ctg ctg tac agc gga gtc cca ttc ctg ggc ccc tac cac Ala Leu Lys Leu Leu Tyr Ser Gly Val Pro Phe Leu Gly Pro Tyr His 125 130 135	436
aag gag tcg gct gtg acg gcc ttc agc gag ggc agc gtc atc gcc tac Lys Glu Ser Ala Val Thr Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr 140 145 150	484
tac tgg tct gag ttc agc atc ccg cag cac ctg gtg gag gag gcc gag Tyr Trp Ser Glu Phe Ser Ile Pro Gln His Leu Val Glu Glu Ala Glu 155 160 165 170	532
cgc gtc atg gcc gag gag cgc gta gtc atg ctg ccc ccg cgg gcg cgc Arg Val Met Ala Glu Glu Arg Val Val Met Leu Pro Pro Arg Ala Arg 175 180 185	580
tcc ctg aag tcc ttt gtg gtc acc tca gtg gtg gct ttc ccc acg gac Ser Leu Lys Ser Phe Val Val Thr Ser Val Val Ala Phe Pro Thr Asp 190 195 200	628
tcc aaa aca gta cag agg acc cag gac aac agc tgc agc ttt ggc ctg Ser Lys Thr Val Gln Arg Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu 205 210 215	676
cac gcc cgc ggt gtg gag ctg atg cgc ttc acc acg ccc ggc ttc cct His Ala Arg Gly Val Glu Leu Met Arg Phe Thr Thr Pro Gly Phe Pro 220 225 230	724
gac agc ccc tac ccc gct cat gcc cgc tgc cag tgg gcc ctg cgg ggg Asp Ser Pro Tyr Pro Ala His Ala Arg Cys Gln Trp Ala Leu Arg Gly 235 240 245 250	772
gac gcc gac tca gtg ctg agc ctc acc ttc cgc agc ttt gac ctt gcg Asp Ala Asp Ser Val Leu Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala 255 260 265	820
tcc tgc gac gag cgc ggc agc gac ctg gtg acg gtg tac aac acc ctg Ser Cys Asp Glu Arg Gly Ser Asp Leu Val Thr Val Tyr Asn Thr Leu 270 275 280	868
agc ccc atg gag ccc cac gcc ctg gtg cag ttg tgt ggc acc tac cct Ser Pro Met Glu Pro His Ala Leu Val Gln Leu Cys Gly Thr Tyr Pro 285 290 295	916
ccc tcc tac aac ctg acc ttc cac tcc tcc cag aac gtc ctg ctc atc Pro Ser Tyr Asn Leu Thr Phe His Ser Ser Gln Asn Val Leu Leu Ile 300 305 310	964
aca ctg ata acc aac act gag cgg cgg cat ccc ggc ttt gag gcc acc Thr Leu Ile Thr Asn Thr Glu Arg Arg His Pro Gly Phe Glu Ala Thr 315 320 325 330	1012
ttc ttc cag ctg cct agg atg agc agc tgt gga ggc cgc tta cgt aaa	1060

-3-

Phe	Phe	Gln	Leu	Pro	Arg	Met	Ser	Ser	Cys	Gly	Gly	Arg	Leu	Arg	Lys	
				335					340					345		
gcc	cag	ggg	aca	ttc	aac	agc	ccc	tac	tac	cca	ggc	cac	tac	cca	ccc	1108
Ala	Gln	Gly	Thr	Phe	Asn	Ser	Pro	Tyr	Tyr	Pro	Gly	His	Tyr	Pro	Pro	
			350					355					360			
aac	att	gac	tgc	aca	tgg	aac	att	gag	gtg	ccc	aac	aac	cag	cat	gtg	1156
Asn	Ile	Asp	Cys	Thr	Trp	Asn	Ile	Glu	Val	Pro	Asn	Asn	Gln	His	Val	
		365					370					375				
aag	gtg	agc	ttc	aaa	ttc	ttc	tac	ctg	ctg	gag	ccc	ggc	gtg	cct	gcg	1204
Lys	Val	Ser	Phe	Lys	Phe	Phe	Tyr	Leu	Leu	Glu	Pro	Gly	Val	Pro	Ala	
	380					385					390					
ggc	acc	tgc	ccc	aag	gac	tac	gtg	gag	atc	aac	ggg	gag	aaa	tac	tgc	1252
Gly	Thr	Cys	Pro	Lys	Asp	Tyr	Val	Glu	Ile	Asn	Gly	Glu	Lys	Tyr	Cys	
395				400						405					410	
gga	gag	agg	tcc	cag	ttc	gtc	gtc	acc	agc	aac	agc	aac	aag	atc	aca	1300
Gly	Glu	Arg	Ser	Gln	Phe	Val	Val	Thr	Ser	Asn	Ser	Asn	Lys	Ile	Thr	
				415				420						425		
gtt	cgc	ttc	cac	tca	gat	cag	tcc	tac	acc	gac	acc	ggc	ttc	tta	gct	1348
Val	Arg	Phe	His	Ser	Asp	Gln	Ser	Tyr	Thr	Asp	Thr	Gly	Phe	Leu	Ala	
			430					435					440			
gaa	tac	ctc	tcc	tac	gac	tcc	agt	gac	cca	tgc	ccg	ggg	cag	ttc	acg	1396
Glu	Tyr	Leu	Ser	Tyr	Asp	Ser	Ser	Asp	Pro	Cys	Pro	Gly	Gln	Phe	Thr	
		445					450					455				
tgc	cgc	acg	ggg	cgg	tgt	atc	cgg	aag	gag	ctg	cgc	tgt	gat	ggc	tgg	1444
Cys	Arg	Thr	Gly	Arg	Cys	Ile	Arg	Lys	Glu	Leu	Arg	Cys	Asp	Gly	Trp	
	460					465					470					
gcc	gac	tgc	acc	gac	cac	agc	gat	gag	ctc	aac	tgc	agt	tgc	gac	gcc	1492
Ala	Asp	Cys	Thr	Asp	His	Ser	Asp	Glu	Leu	Asn	Cys	Ser	Cys	Asp	Ala	
475					480					485					490	
ggc	cac	cag	ttc	acg	tgc	aag	aac	aag	ttc	tgc	aag	ccc	ctc	ttc	tgg	1540
Gly	His	Gln	Phe	Thr	Cys	Lys	Asn	Lys	Phe	Cys	Lys	Pro	Leu	Phe	Trp	
			495					500						505		
gtc	tgc	gac	agt	gtg	aac	gac	tgc	gga	gac	aac	agc	gac	gag	cag	ggg	1588
Val	Cys	Asp	Ser	Val	Asn	Asp	Cys	Gly	Asp	Asn	Ser	Asp	Glu	Gln	Gly	
			510					515					520			
tgc	agt	tgt	ccg	gcc	cag	acc	ttc	agg	tgt	tcc	aac	ggg	aag	tgc	ctc	1636
Cys	Ser	Cys	Pro	Ala	Gln	Thr	Phe	Arg	Cys	Ser	Asn	Gly	Lys	Cys	Leu	
		525					530					535				
tcg	aaa	agc	cag	cag	tgc	aac	ggg	aag	gac	gac	tgt	ggg	gac	ggg	tcc	1684
Ser	Lys	Ser	Gln	Gln	Cys	Asn	Gly	Lys	Asp	Asp	Cys	Gly	Asp	Gly	Ser	
	540					545					550					
gac	gag	gcc	tcc	tgc	ccc	aag	gtg	aac	gtc	gtc	act	tgt	acc	aaa	cac	1732
Asp	Glu	Ala	Ser	Cys	Pro	Lys	Val	Asn	Val	Val	Thr	Cys	Thr	Lys	His	
555					560					565					570	
acc	tac	cgc	tgc	ctc	aac	ggg	ctc	tgc	ttg	agc	aag	ggc	aac	cct	gag	1780
Thr	Tyr	Arg	Cys	Leu	Asn	Gly	Leu	Cys	Leu	Ser	Lys	Gly	Asn	Pro	Glu	

-4-

575								580				585				
tgt	gac	ggg	aag	gag	gac	tgt	agc	gac	ggc	tca	gat	gag	aag	gac	tgc	1828
Cys	Asp	Gly	Lys	Glu	Asp	Cys	Ser	Asp	Gly	Ser	Asp	Glu	Lys	Asp	Cys	
			590					595					600			
gac	tgt	ggg	ctg	cgg	tca	ttc	acg	aga	cag	gct	cgt	gtt	gtt	ggg	ggc	1876
Asp	Cys	Gly	Leu	Arg	Ser	Phe	Thr	Arg	Gln	Ala	Arg	Val	Val	Gly	Gly	
		605					610					615				
acg	gat	gcg	gat	gag	ggc	gag	tgg	ccc	tgg	cag	gta	agc	ctg	cat	gct	1924
Thr	Asp	Ala	Asp	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Val	Ser	Leu	His	Ala	
	620					625					630					
ctg	ggc	cag	ggc	cac	atc	tgc	ggt	gct	tcc	ctc	atc	tct	ccc	aac	tgg	1972
Leu	Gly	Gln	Gly	His	Ile	Cys	Gly	Ala	Ser	Leu	Ile	Ser	Pro	Asn	Trp	
	635				640					645					650	
ctg	gtc	tct	gcc	gca	cac	tgc	tac	atc	gat	gac	aga	gga	ttc	agg	tac	2020
Leu	Val	Ser	Ala	Ala	His	Cys	Tyr	Ile	Asp	Asp	Arg	Gly	Phe	Arg	Tyr	
			655						660					665		
tca	gac	ccc	acg	cag	tgg	acg	gcc	ttc	ctg	ggc	ttg	cac	gac	cag	agc	2068
Ser	Asp	Pro	Thr	Gln	Trp	Thr	Ala	Phe	Leu	Gly	Leu	His	Asp	Gln	Ser	
			670					675					680			
cag	cgc	agc	gcc	cct	ggg	gtg	cag	gag	cgc	agg	ctc	aag	cgc	atc	atc	2116
Gln	Arg	Ser	Ala	Pro	Gly	Val	Gln	Glu	Arg	Arg	Leu	Lys	Arg	Ile	Ile	
		685				690						695				
tcc	cac	ccc	ttc	ttc	aat	gac	ttc	acc	ttc	gac	tat	gac	atc	gcg	ctg	2164
Ser	His	Pro	Phe	Phe	Asn	Asp	Phe	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	
	700					705					710					
ctg	gag	ctg	gag	aaa	ccg	gca	gag	tac	agc	tcc	atg	gtg	cgg	ccc	atc	2212
Leu	Glu	Leu	Glu	Lys	Pro	Ala	Glu	Tyr	Ser	Ser	Met	Val	Arg	Pro	Ile	
	715				720					725					730	
tgc	ctg	ccg	gac	gcc	tcc	cat	gtc	ttc	cct	gcc	ggc	aag	gcc	atc	tgg	2260
Cys	Leu	Pro	Asp	Ala	Ser	His	Val	Phe	Pro	Ala	Gly	Lys	Ala	Ile	Trp	
			735						740					745		
gtc	acg	ggc	tgg	gga	cac	acc	cag	tat	gga	ggc	act	ggc	gcg	ctg	atc	2308
Val	Thr	Gly	Trp	Gly	His	Thr	Gln	Tyr	Gly	Gly	Thr	Gly	Ala	Leu	Ile	
		750					755						760			
ctg	caa	aag	ggt	gag	atc	cgc	gtc	atc	aac	cag	acc	acc	tgc	gag	aac	2356
Leu	Gln	Lys	Gly	Glu	Ile	Arg	Val	Ile	Asn	Gln	Thr	Thr	Cys	Glu	Asn	
		765					770					775				
ctc	ctg	ccg	cag	cag	atc	acg	ccg	cgc	atg	atg	tgc	gtg	ggc	ttc	ctc	2404
Leu	Leu	Pro	Gln	Gln	Ile	Thr	Pro	Arg	Met	Met	Cys	Val	Gly	Phe	Leu	
	780					785					790					
agc	ggc	ggc	gtg	gac	tcc	tgc	cag	ggt	gat	tcc	ggg	gga	ccc	ctg	tcc	2452
Ser	Gly	Gly	Val	Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ser	
	795				800					805					810	
agc	gtg	gag	gcg	gat	ggg	cgg	atc	ttc	cag	gcc	ggt	gtg	gtg	agc	tgg	2500
Ser	Val	Glu	Ala	Asp	Gly	Arg	Ile	Phe	Gln	Ala	Gly	Val	Val	Ser	Trp	
			815						820					825		

-5-

gga gac ggc tgc gct cag agg aac aag cca ggc gtg tac aca agg ctc 2548  
 Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu  
                   830                                  835                                  840

cct ctg ttt cgg gac tgg atc aaa gag aac act ggg gta ta ggggcccgggg 2599  
 Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly Val  
                   845                                  850                                  855

ccacccaaat gtgtacacct gcgggggccac ccacgtgcc cccagtggtg cagcctgca 2659  
 ggctggagac tggaccgctg actgcaccag cgccccaga acatacactg tgaactcaat 2719  
 ctccagggct ccaaattctgc ctagaaaacc tctcgttcc tcagcctcca aagtggagct 2779  
 gggaggtaga aggggaggac actggtggtt ctactgaccc aactgggggc aaagggttga 2839  
 agacacagcc tcccccgcca gcccgaagct gggccgaggc gcgtttgtgt atatctgcct 2899  
 cccctgtctg taaggagcag cggggaacgga gcttcggagc ctcctcagtg aagggtggtg 2959  
 ggctgccgga tctgggctgt gggggcccttg ggccacgctc ttgaggaagc ccaggctcgg 3019  
 aggaccctgg aaaacagacg ggtctgagac tgaaattgtt ttaccagctc ccagggtgga 3079  
 cttcagtggtg tgtatttctg taaatgggta aaacaattta tttcttttta aaaaaaaaaa 3139  
 aaaaaaaaaa 3147

<210> 2  
 <211> 855  
 <212> PRT  
 <213> Homo Sapien

<400> 2  
 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Gly Pro Lys Asp Phe  
   1                  5                  10                  15  
 Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn Gly Leu  
                   20                  25                  30  
 Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys Lys Val Glu  
                   35                  40                  45  
 Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala Val Leu Ile Gly  
                   50                  55                  60  
 Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln  
   65                  70                  75                  80  
 Tyr Arg Asp Val Arg Val Gln Lys Val Phe Asn Gly Tyr Met Arg Ile  
                   85                  90                  95  
 Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu Asn Ser Asn Ser Thr Glu  
                   100                  105                  110  
 Phe Val Ser Leu Ala Ser Lys Val Lys Asp Ala Leu Lys Leu Leu Tyr  
                   115                  120                  125  
 Ser Gly Val Pro Phe Leu Gly Pro Tyr His Lys Glu Ser Ala Val Thr  
                   130                  135                  140  
 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser  
   145                  150                  155                  160  
 Ile Pro Gln His Leu Val Glu Glu Ala Glu Arg Val Met Ala Glu Glu  
                   165                  170                  175  
 Arg Val Val Met Leu Pro Pro Arg Ala Arg Ser Leu Lys Ser Phe Val  
                   180                  185                  190  
 Val Thr Ser Val Val Ala Phe Pro Thr Asp Ser Lys Thr Val Gln Arg  
                   195                  200                  205  
 Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu His Ala Arg Gly Val Glu  
   210                  215                  220  
 Leu Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala  
   225                  230                  235                  240  
 His Ala Arg Cys Gln Trp Ala Leu Arg Gly Asp Ala Asp Ser Val Leu  
                   245                  250                  255  
 Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala Ser Cys Asp Glu Arg Gly  
                   260                  265                  270  
 Ser Asp Leu Val Thr Val Tyr Asn Thr Leu Ser Pro Met Glu Pro His



-6-

Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	Pro	Ser	Tyr	Asn	Leu	Thr
290						295					300				
Phe	His	Ser	Ser	Gln	Asn	Val	Leu	Leu	Ile	Thr	Leu	Ile	Thr	Asn	Thr
305					310					315					320
Glu	Arg	Arg	His	Pro	Gly	Phe	Glu	Ala	Thr	Phe	Phe	Gln	Leu	Pro	Arg
				325					330					335	
Met	Ser	Ser	Cys	Gly	Gly	Arg	Leu	Arg	Lys	Ala	Gln	Gly	Thr	Phe	Asn
			340					345					350		
Ser	Pro	Tyr	Tyr	Pro	Gly	His	Tyr	Pro	Pro	Asn	Ile	Asp	Cys	Thr	Trp
		355					360					365			
Asn	Ile	Glu	Val	Pro	Asn	Asn	Gln	His	Val	Lys	Val	Ser	Phe	Lys	Phe
370						375					380				
Phe	Tyr	Leu	Leu	Glu	Pro	Gly	Val	Pro	Ala	Gly	Thr	Cys	Pro	Lys	Asp
385					390					395					400
Tyr	Val	Glu	Ile	Asn	Gly	Glu	Lys	Tyr	Cys	Gly	Glu	Arg	Ser	Gln	Phe
				405					410					415	
Val	Val	Thr	Ser	Asn	Ser	Asn	Lys	Ile	Thr	Val	Arg	Phe	His	Ser	Asp
			420					425					430		
Gln	Ser	Tyr	Thr	Asp	Thr	Gly	Phe	Leu	Ala	Glu	Tyr	Leu	Ser	Tyr	Asp
		435					440					445			
Ser	Ser	Asp	Pro	Cys	Pro	Gly	Gln	Phe	Thr	Cys	Arg	Thr	Gly	Arg	Cys
		450				455					460				
Ile	Arg	Lys	Glu	Leu	Arg	Cys	Asp	Gly	Trp	Ala	Asp	Cys	Thr	Asp	His
465					470					475					480
Ser	Asp	Glu	Leu	Asn	Cys	Ser	Cys	Asp	Ala	Gly	His	Gln	Phe	Thr	Cys
				485					490					495	
Lys	Asn	Lys	Phe	Cys	Lys	Pro	Leu	Phe	Trp	Val	Cys	Asp	Ser	Val	Asn
			500					505					510		
Asp	Cys	Gly	Asp	Asn	Ser	Asp	Glu	Gln	Gly	Cys	Ser	Cys	Pro	Ala	Gln
		515					520					525			
Thr	Phe	Arg	Cys	Ser	Asn	Gly	Lys	Cys	Leu	Ser	Lys	Ser	Gln	Gln	Cys
		530				535					540				
Asn	Gly	Lys	Asp	Asp	Cys	Gly	Asp	Gly	Ser	Asp	Glu	Ala	Ser	Cys	Pro
545					550					555					560
Lys	Val	Asn	Val	Val	Thr	Cys	Thr	Lys	His	Thr	Tyr	Arg	Cys	Leu	Asn
				565					570					575	
Gly	Leu	Cys	Leu	Ser	Lys	Gly	Asn	Pro	Glu	Cys	Asp	Gly	Lys	Glu	Asp
			580					585					590		
Cys	Ser	Asp	Gly	Ser	Asp	Glu	Lys	Asp	Cys	Asp	Cys	Gly	Leu	Arg	Ser
		595				600						605			
Phe	Thr	Arg	Gln	Ala	Arg	Val	Val	Gly	Gly	Thr	Asp	Ala	Asp	Glu	Gly
		610				615					620				
Glu	Trp	Pro	Trp	Gln	Val	Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile
625					630					635					640
Cys	Gly	Ala	Ser	Leu	Ile	Ser	Pro	Asn	Trp	Leu	Val	Ser	Ala	Ala	His
				645					650					655	
Cys	Tyr	Ile	Asp	Asp	Arg	Gly	Phe	Arg	Tyr	Ser	Asp	Pro	Thr	Gln	Trp
			660					665					670		
Thr	Ala	Phe	Leu	Gly	Leu	His	Asp	Gln	Ser	Gln	Arg	Ser	Ala	Pro	Gly
		675					680						685		
Val	Gln	Glu	Arg	Arg	Leu	Lys	Arg	Ile	Ile	Ser	His	Pro	Phe	Phe	Asn
		690				695					700				
Asp	Phe	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Glu	Leu	Glu	Lys	Pro
705					710					715					720
Ala	Glu	Tyr	Ser	Ser	Met	Val	Arg	Pro	Ile	Cys	Leu	Pro	Asp	Ala	Ser
				725					730					735	
His	Val	Phe	Pro	Ala	Gly	Lys	Ala	Ile	Trp	Val	Thr	Gly	Trp	Gly	His
			740					745					750		
Thr	Gln	Tyr	Gly	Gly	Thr	Gly	Ala	Leu	Ile	Leu	Gln	Lys	Gly	Glu	Ile
		755					760					765			

-7-

Arg Val Ile Asn Gln Thr Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile  
 770 775 780  
 Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser  
 785 790 795 800  
 Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Val Glu Ala Asp Gly  
 805 810 815  
 Arg Ile Phe Gln Ala Gly Val Val Ser Trp Gly Asp Gly Cys Ala Gln  
 820 825 830  
 Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp  
 835 840 845  
 Ile Lys Glu Asn Thr Gly Val  
 850 855

<210> 3  
 <211> 2137  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <221> CDS  
 <222> (261)...(1574)  
 <223> Nucleic acid encoding a transmembrane serine  
 protease (MTSP3) protein

<400> 3  
 ccataccta acgactcact atagggctcg agcggccgcc cgggcagggtc agagagagggc 60  
 agcagcttgc tcagcggaca aggatgctgg gcgtgaggga ccaaggcctg ccctgcactc 120  
 gggcctcctc cagccagtgc tgaccaggga cttctgacct gctggccagc caggacctgt 180  
 gtggggaggc cctcctgctg ccttgggggtg acaatctcag ctccaggcta caggagagacc 240  
 gggaggatca cagagccagc atg tta cag gat cct gac agt gat caa cct ctg 293  
 Met Leu Gln Asp Pro Asp Ser Asp Gln Pro Leu  
 1 5 10  
 aac agc ctc gat gtc aaa ccc ctg cgc aaa ccc cgt atc ccc atg gag 341  
 Asn Ser Leu Asp Val Lys Pro Leu Arg Lys Pro Arg Ile Pro Met Glu  
 15 20 25  
 acc ttc aga aag gtg ggg atc ccc atc atc ata gca cta ctg agc ctg 389  
 Thr Phe Arg Lys Val Gly Ile Pro Ile Ile Ile Ala Leu Leu Ser Leu  
 30 35 40  
 gcg agt atc atc att gtg gtt gtc ctc atc aag gtg att ctg gat aaa 437  
 Ala Ser Ile Ile Ile Val Val Val Leu Ile Lys Val Ile Leu Asp Lys  
 45 50 55  
 tac tac ttc ctc tgc ggg cag cct ctc cac ttc atc ccg agg aag cag 485  
 Tyr Tyr Phe Leu Cys Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln  
 60 65 70 75  
 ctg tgt gac gga gag ctg gac tgt ccc ttg ggg gag gac gag gag cac 533  
 Leu Cys Asp Gly Glu Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu His  
 80 85 90  
 tgt gtc aag agc ttc ccc gaa ggg cct gca gtg gca gtc cgc ctc tcc 581  
 Cys Val Lys Ser Phe Pro Glu Gly Pro Ala Val Ala Val Arg Leu Ser  
 95 100 105  
 aag gac cga tcc aca ctg cag gtg ctg gac tcg gcc aca ggg aac tgg 629  
 Lys Asp Arg Ser Thr Leu Gln Val Leu Asp Ser Ala Thr Gly Asn Trp  
 110 115 120

-8-

ttc tct gcc tgt ttc gac aac ttc aca gaa gct ctc gct gag aca gcc Phe Ser Ala Cys Phe Asp Asn Phe Thr Glu Ala Leu Ala Glu Thr Ala 125 130 135	677
tgt agg cag atg ggc tac agc agc aaa ccc acc ttc aga gct gtg gag Cys Arg Gln Met Gly Tyr Ser Ser Lys Pro Thr Phe Arg Ala Val Glu 140 145 150 155	725
att ggc cca gac cag gat ctg gat gtt gtt gaa atc aca gaa aac agc Ile Gly Pro Asp Gln Asp Leu Asp Val Val Glu Ile Thr Glu Asn Ser 160 165 170	773
cag gag ctt cgc atg cgg aac tca agt ggg ccc tgt ctc tca ggc tcc Gln Glu Leu Arg Met Arg Asn Ser Ser Gly Pro Cys Leu Ser Gly Ser 175 180 185	821
ctg gtc tcc ctg cac tgt ctt gcc tgt ggg aag agc ctg aag acc ccc Leu Val Ser Leu His Cys Leu Ala Cys Gly Lys Ser Leu Lys Thr Pro 190 195 200	869
cgt gtg gtg ggt ggg gag gag gcc tct gtg gat tct tgg cct tgg cag Arg Val Val Gly Gly Glu Glu Ala Ser Val Asp Ser Trp Pro Trp Gln 205 210 215	917
gtc agc atc cag tac gac ata cag cac gtc tgt gga ggg agc atc ctg Val Ser Ile Gln Tyr Asp Ile Gln His Val Cys Gly Gly Ser Ile Leu 220 225 230 235	965
gac ccc cac tgg gtc ctc acg gca gcc cac tgc ttc agg aaa cat acc Asp Pro His Trp Val Leu Thr Ala Ala His Cys Phe Arg Lys His Thr 240 245 250	1013
gat gtg ttc aac tgg aag gtg cgg gca ggc tca gac aaa ctg ggc agc Asp Val Phe Asn Trp Lys Val Arg Ala Gly Ser Asp Lys Leu Gly Ser 255 260 265	1061
ttc cca tcc ctg gct gtg gcc aag atc atc atc att gaa ttc aac ccc Phe Pro Ser Leu Ala Val Ala Lys Ile Ile Ile Ile Glu Phe Asn Pro 270 275 280	1109
atg tac ccc aaa gac aat gac atc gcc ctc atg aag ctg cag ttc cca Met Tyr Pro Lys Asp Asn Asp Ile Ala Leu Met Lys Leu Gln Phe Pro 285 290 295	1157
ctc act ttc tca ggc aca gtc agg ctc atc tgt ctg ccc ttc ttt gat Leu Thr Phe Ser Gly Thr Val Arg Leu Ile Cys Leu Pro Phe Phe Asp 300 305 310 315	1205
gag gag ctc act cca gcc acc cca ctc tgg atc att gga tgg ggc ttt Glu Glu Leu Thr Pro Ala Thr Pro Leu Trp Ile Ile Gly Trp Gly Phe 320 325 330	1253
acg aag cag aat gga ggg aag atg tct gac ata ctg ctg cag gcg tca Thr Lys Gln Asn Gly Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser 335 340 345	1301
gtc cag gtc att gac agc aca cgg tgc aat gca gac gat gcg tac cag Val Gln Val Ile Asp Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln 350 355 360	1349
ggg gaa gtc acc gag aag atg atg tgt gca ggc atc ccg gaa ggg ggt	1397

-9-

Gly Glu Val Thr Glu Lys Met Met Cys Ala Gly Ile Pro Glu Gly Gly  
 365 370 375  
 gtg gac acc tgc cag ggt gac agt ggt ggg ccc ctg atg tac caa tct 1445  
 Val Asp Thr Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser 395  
 380 385 390  
 gac cag tgg cat gtg gtg ggc atc gtt agc tgg ggc tat ggc tgc ggg 1493  
 Asp Gln Trp His Val Val Gly Ile Val Ser Trp Gly Tyr Gly Cys Gly 410  
 400 405  
 ggc ccg agc acc cca gga gta tac acc aag gtc tca gcc tat ctc aac 1541  
 Gly Pro Ser Thr Pro Gly Val Tyr Thr Lys Val Ser Ala Tyr Leu Asn 425  
 415 420  
 tgg atc tac aat gtc tgg aag gct gag ctg taa tgctgctgcc cctttgcagt 1594  
 Trp Ile Tyr Asn Val Trp Lys Ala Glu Leu \* 435  
 430  
 gctgggagcc gcttccttcc tgccctgccc acctggggat cccccaagat cagacacaga 1654  
 gcaagagtcc ccttgggtac acccctctgc ccacagcctc agcatttctt ggagcagcaa 1714  
 agggcctcaa ttcctgtaag agaccctcgc agcccagagg cgcccagagg aagtcagcag 1774  
 ccctagctcg gccacacttg gtgctccag catcccaggg agagacacag cccactgaac 1834  
 aaggtctcag gggatttgct aagccaagaa ggaactttcc cacactactg aatggaagca 1894  
 ggctgtcttg taaaagccca gatcactgtg ggctggagag gagaaggaaa gggctctgcgc 1954  
 cagccctgtc cgtcttcacc catccccaag cctactagag caagaaacca gttgtaatat 2014  
 aaaatgcact gccctactgt tggatgact accgttacct actgttgtca ttgttattac 2074  
 agctatggcc actattatta aagagctgtg taacaaaaaa aaaaaaaaaa aaaaaaaaaa 2134  
 aaa 2137  
 <210> 4  
 <211> 437  
 <212> PRT  
 <213> Homo Sapien  
 <400> 4  
 Met Leu Gln Asp Pro Asp Ser Asp Gln Pro Leu Asn Ser Leu Asp Val  
 1 5 10 15  
 Lys Pro Leu Arg Lys Pro Arg Ile Pro Met Glu Thr Phe Arg Lys Val  
 20 25 30  
 Gly Ile Pro Ile Ile Ile Ala Leu Ser Leu Ala Ser Ile Ile Ile  
 35 40 45  
 Val Val Val Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Leu Cys  
 50 55 60  
 Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp Gly Glu  
 65 70 75 80  
 Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu His Cys Val Lys Ser Phe  
 85 90 95  
 Pro Glu Gly Pro Ala Val Ala Val Arg Leu Ser Lys Asp Arg Ser Thr  
 100 105 110  
 Leu Gln Val Leu Asp Ser Ala Thr Gly Asn Trp Phe Ser Ala Cys Phe  
 115 120 125  
 Asp Asn Phe Thr Glu Ala Leu Ala Glu Thr Ala Cys Arg Gln Met Gly  
 130 135 140  
 Tyr Ser Ser Lys Pro Thr Phe Arg Ala Val Glu Ile Gly Pro Asp Gln  
 145 150 155 160  
 Asp Leu Asp Val Val Glu Ile Thr Glu Asn Ser Gln Glu Leu Arg Met  
 165 170 175  
 Arg Asn Ser Ser Gly Pro Cys Leu Ser Gly Ser Leu Val Ser Leu His  
 180 185 190  
 Cys Leu Ala Cys Gly Lys Ser Leu Lys Thr Pro Arg Val Val Gly Gly

-10-

[illegible]

```
<210> 5
<211> 708
<212> DNA
<213> Homo Sapien
```

```
<220>
<221> CDS
<222> (1)...(708)
<223> Nucleic acid encoding an MTSP4 protease domain
```

<400> 5																
att	gtt	ggt	gga	gct	gtg	tcc	tcc	gag	ggt	gag	tgg	cca	tgg	cag	gcc	48
Ile	Val	Gly	Gly	Ala	Val	Ser	Ser	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Ala	
1				5					10					15		
agc	ctc	cag	gtt	cgg	ggt	cga	cac	atc	tgt	ggg	ggg	gcc	ctc	atc	gct	96
Ser	Leu	Gln	Val	Arg	Gly	Arg	His	Ile	Cys	Gly	Gly	Ala	Leu	Ile	Ala	
			20					25					30			
gac	cgc	tgg	gtg	ata	aca	gct	gcc	cac	tgc	ttc	cag	gag	gac	agc	atg	144
Asp	Arg	Trp	Val	Ile	Thr	Ala	Ala	His	Cys	Phe	Gln	Glu	Asp	Ser	Met	
		35				40						45				
gcc	tcc	acg	gtg	ctg	tgg	acc	gtg	ttc	ctg	ggc	aag	gtg	tgg	cag	aac	192
Ala	Ser	Thr	Val	Leu	Trp	Thr	Val	Phe	Leu	Gly	Lys	Val	Trp	Gln	Asn	
	50					55					60					
tcg	cgc	tgg	cct	gga	gag	gtg	tcc	ttc	aag	gtg	agc	cgc	ctg	ctc	ctg	240
Ser	Arg	Trp	Pro	Gly	Glu	Val	Ser	Phe	Lys	Val	Ser	Arg	Leu	Leu	Leu	

-11-

65	70	75	80	
cac ccg tac cac gaa gag gac agc cat gac tac gac gtg gcg ctg ctg				288
His Pro Tyr His Glu Glu Asp Ser His Asp Tyr Asp Val Ala Leu Leu	85	90	95	
cag ctc gac cac ccg gtg gtg cgc tcg gcc gcc gtg cgc ccc gtc tgc				336
Gln Leu Asp His Pro Val Val Arg Ser Ala Ala Val Arg Pro Val Cys	100	105	110	
ctg ccc gcg cgc tcc cac ttc ttc gag ccc gcc ctg cac tgc tgg att				384
Leu Pro Ala Arg Ser His Phe Phe Glu Pro Gly Leu His Cys Trp Ile	115	120	125	
acg gcc tgg gcc gcc ttg cgc gag gcc gcc ccc atc agc aac gct ctg				432
Thr Gly Trp Gly Ala Leu Arg Glu Gly Gly Pro Ile Ser Asn Ala Leu	130	135	140	
cag aaa gtg gat gtg cag ttg atc cca cag gac ctg tgc agc gag gtc				480
Gln Lys Val Asp Val Gln Leu Ile Pro Gln Asp Leu Cys Ser Glu Val	145	150	155	160
tat cgc tac cag gtg acg cca cgc atg ctg tgt gcc gcc tac cgc aag				528
Tyr Arg Tyr Gln Val Thr Pro Arg Met Leu Cys Ala Gly Tyr Arg Lys	165	170	175	
ggc aag aag gat gcc tgt cag ggt gac tca ggt ggt ccg ctg gtg tgc				576
Gly Lys Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys	180	185	190	
aag gca ctc agt gcc cgc tgg ttc ctg gcg ggg ctg gtc agc tgg gcc				624
Lys Ala Leu Ser Gly Arg Trp Phe Leu Ala Gly Leu Val Ser Trp Gly	195	200	205	
ctg gcc tgt gcc cgg cct aac tac ttc gcc gtc tac acc cgc atc aca				672
Leu Gly Cys Gly Arg Pro Asn Tyr Phe Gly Val Tyr Thr Arg Ile Thr	210	215	220	
ggt gtg atc agc tgg atc cag caa gtg gtg acc tga				708
Gly Val Ile Ser Trp Ile Gln Gln Val Val Thr *	225	230	235	

<210> 6  
 <211> 235  
 <212> PRT  
 <213> Homo Sapien

<400> 6  
 Ile Val Gly Gly Ala Val Ser Ser Glu Gly Glu Trp Pro Trp Gln Ala  
 1 5 10 15  
 Ser Leu Gln Val Arg Gly Arg His Ile Cys Gly Gly Ala Leu Ile Ala  
 20 25 30  
 Asp Arg Trp Val Ile Thr Ala Ala His Cys Phe Gln Glu Asp Ser Met  
 35 40 45  
 Ala Ser Thr Val Leu Trp Thr Val Phe Leu Gly Lys Val Trp Gln Asn  
 50 55 60  
 Ser Arg Trp Pro Gly Glu Val Ser Phe Lys Val Ser Arg Leu Leu Leu  
 65 70 75 80  
 His Pro Tyr His Glu Glu Asp Ser His Asp Tyr Asp Val Ala Leu Leu  
 85 90 95

-12-

Gln Leu Asp His Pro Val Val Arg Ser Ala Ala Val Arg Pro Val Cys  
 100 105 110  
 Leu Pro Ala Arg Ser His Phe Phe Glu Pro Gly Leu His Cys Trp Ile  
 115 120 125  
 Thr Gly Trp Gly Ala Leu Arg Glu Gly Gly Pro Ile Ser Asn Ala Leu  
 130 135 140  
 Gln Lys Val Asp Val Gln Leu Ile Pro Gln Asp Leu Cys Ser Glu Val  
 145 150 155 160  
 Tyr Arg Tyr Gln Val Thr Pro Arg Met Leu Cys Ala Gly Tyr Arg Lys  
 165 170 175  
 Gly Lys Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys  
 180 185 190  
 Lys Ala Leu Ser Gly Arg Trp Phe Leu Ala Gly Leu Val Ser Trp Gly  
 195 200 205  
 Leu Gly Cys Gly Arg Pro Asn Tyr Phe Gly Val Tyr Thr Arg Ile Thr  
 210 215 220 225  
 Gly Val Ile Ser Trp Ile Gln Gln Val Val Thr  
 230 235

<210> 7  
 <211> 3104  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <221> CDS  
 <222> (33)...(2441)  
 <223> Nucleic acid encoding MTSP4-L (long form) splice variant

<400> 7  
 tcatcggcca gaggggtgatc agtgagcaga ag atg ccc gtg gcc gag gcc ccc 53  
 Met Pro Val Ala Glu Ala Pro  
 1 5  
 cag gtg gct ggc ggg cag ggg gac gga ggt gat ggc gag gaa gcg gag 101  
 Gln Val Ala Gly Gly Gln Gly Asp Gly Gly Asp Gly Glu Glu Ala Glu  
 10 15 20  
 ccg gag ggg atg ttc aag gcc tgt gag gac tcc aag aga aaa gcc cgg 149  
 Pro Glu Gly Met Phe Lys Ala Cys Glu Asp Ser Lys Arg Lys Ala Arg  
 25 30 35  
 ggc tac ctc cgc ctg gtg ccc ctg ttt gtg ctg ctg gcc ctg ctc gtg 197  
 Gly Tyr Leu Arg Leu Val Pro Leu Phe Val Leu Leu Ala Leu Leu Val  
 40 45 50 55  
 ctg gct tcg gcg ggg gtg cta ctc tgg tat ttc cta ggg tac aag gcg 245  
 Leu Ala Ser Ala Gly Val Leu Leu Trp Tyr Phe Leu Gly Tyr Lys Ala  
 60 65 70  
 gag gtg atg gtc agc cag gtg tac tca ggc agt ctg cgt gta ctc aat 293  
 Glu Val Met Val Ser Gln Val Tyr Ser Gly Ser Leu Arg Val Leu Asn  
 75 80 85  
 cgc cac ttc tcc cag gat ctt acc cgc cgg gaa tct agt gcc ttc cgc 341  
 Arg His Phe Ser Gln Asp Leu Thr Arg Arg Glu Ser Ser Ala Phe Arg  
 90 95 100  
 agt gaa acc gcc aaa gcc cag aag atg ctc aag gag ctc atc acc agc 389  
 Ser Glu Thr Ala Lys Ala Gln Lys Met Leu Lys Glu Leu Ile Thr Ser  
 105 110 115

-13-

acc	cgc	ctg	gga	act	tac	tac	aac	tcc	agc	tcc	gtc	tat	tcc	ttt	ggg	437
Thr	Arg	Leu	Gly	Thr	Tyr	Tyr	Asn	Ser	Ser	Ser	Val	Tyr	Ser	Phe	Gly	
120					125					130					135	
gag	gga	ccc	ctc	acc	tgc	ttc	ttc	tgg	ttc	att	ctc	caa	atc	ccc	gag	485
Glu	Gly	Pro	Leu	Thr	Cys	Phe	Phe	Trp	Phe	Ile	Leu	Gln	Ile	Pro	Glu	
				140					145					150		
cac	cgc	cgg	ctg	atg	ctg	agc	ccc	gag	gtg	gtg	cag	gca	ctg	ctg	gtg	533
His	Arg	Arg	Leu	Met	Leu	Ser	Pro	Glu	Val	Val	Gln	Ala	Leu	Leu	Val	
			155					160					165			
gag	gag	ctg	ctg	tcc	aca	gtc	aac	agc	tcg	gct	gcc	gtc	ccc	tac	agg	581
Glu	Glu	Leu	Leu	Ser	Thr	Val	Asn	Ser	Ser	Ala	Ala	Val	Pro	Tyr	Arg	
		170					175					180				
gcc	gag	tac	gaa	gtg	gac	ccc	gag	ggc	cta	gtg	atc	ctg	gaa	gcc	agt	629
Ala	Glu	Tyr	Glu	Val	Asp	Pro	Glu	Gly	Leu	Val	Ile	Leu	Glu	Ala	Ser	
	185					190					195					
gtg	aaa	gac	ata	gct	gca	ttg	aac	tcc	acg	ctg	ggc	tgt	tac	cgc	tac	677
Val	Lys	Asp	Ile	Ala	Ala	Leu	Asn	Ser	Thr	Leu	Gly	Cys	Tyr	Arg	Tyr	
200					205					210					215	
agc	tac	gtg	ggc	cag	ggc	cag	gtc	ctc	cgg	ctg	aag	ggg	cct	gac	cac	725
Ser	Tyr	Val	Gly	Gln	Gly	Gln	Val	Leu	Arg	Leu	Lys	Gly	Pro	Asp	His	
				220					225					230		
ctg	gcc	tcc	agc	tgc	ctg	tgg	cac	ctg	cag	ggc	ccc	aag	gac	ctc	atg	773
Leu	Ala	Ser	Ser	Cys	Leu	Trp	His	Leu	Gln	Gly	Pro	Lys	Asp	Leu	Met	
			235					240					245			
ctc	aaa	ctc	cgg	ctg	gag	tgg	acg	ctg	gca	gag	tgc	cgg	gac	cga	ctg	821
Leu	Lys	Leu	Arg	Leu	Glu	Trp	Thr	Leu	Ala	Glu	Cys	Arg	Asp	Arg	Leu	
		250					255					260				
gcc	atg	tat	gac	gtg	gcc	ggg	ccc	ctg	gag	aag	agg	ctc	atc	acc	tcg	869
Ala	Met	Tyr	Asp	Val	Ala	Gly	Pro	Leu	Glu	Lys	Arg	Leu	Ile	Thr	Ser	
	265					270					275					
gtg	tac	ggc	tgc	agc	cgc	cag	gag	ccc	gtg	gtg	gag	gtt	ctg	gcg	tcg	917
Val	Tyr	Gly	Cys	Ser	Arg	Gln	Glu	Pro	Val	Val	Glu	Val	Leu	Ala	Ser	
280					285					290					295	
ggg	gcc	atc	atg	gcg	gtc	gtc	tgg	aag	aag	ggc	ctg	cac	agc	tac	tac	965
Gly	Ala	Ile	Met	Ala	Val	Val	Trp	Lys	Lys	Gly	Leu	His	Ser	Tyr	Tyr	
				300					305					310		
gac	ccc	ttc	gtg	ctc	tcc	gtg	cag	ccg	gtg	gtc	ttc	cag	gcc	tgt	gaa	1013
Asp	Pro	Phe	Val	Leu	Ser	Val	Gln	Pro	Val	Val	Phe	Gln	Ala	Cys	Glu	
			315					320					325			
gtg	aac	ctg	acg	ctg	gac	aac	agg	ctc	gac	tcc	cag	ggc	gtc	ctc	agc	1061
Val	Asn	Leu	Thr	Leu	Asp	Asn	Arg	Leu	Asp	Ser	Gln	Gly	Val	Leu	Ser	
		330					335					340				
acc	ccg	tac	ttc	ccc	agc	tac	tac	tcg	ccc	caa	acc	cac	tgc	tcc	tgg	1109
Thr	Pro	Tyr	Phe	Pro	Ser	Tyr	Tyr	Ser	Pro	Gln	Thr	His	Cys	Ser	Trp	
	345					350					355					



-14-

cac ctc acg gtg ccc tct ctg gac tac ggc ttg gcc ctc tgg ttt gat His Leu Thr Val Pro Ser Leu Asp Tyr Gly Leu Ala Leu Trp Phe Asp 360 365 370 375	1157
gcc tat gca ctg agg agg cag aag tat gat ttg ccg tgc acc cag ggc Ala Tyr Ala Leu Arg Arg Gln Lys Tyr Asp Leu Pro Cys Thr Gln Gly 380 385 390	1205
cag tgg acg atc cag aac agg agg ctg tgt ggc ttg cgc atc ctg cag Gln Trp Thr Ile Gln Asn Arg Arg Leu Cys Gly Leu Arg Ile Leu Gln 395 400 405	1253
ccc tac gcc gag agg atc ccc gtg gtg gcc acg gcc ggg atc acc atc Pro Tyr Ala Glu Arg Ile Pro Val Val Ala Thr Ala Gly Ile Thr Ile 410 415 420	1301
aac ttc acc tcc cag atc tcc ctc acc ggg ccc ggt gtg cgg gtg cac Asn Phe Thr Ser Gln Ile Ser Leu Thr Gly Pro Gly Val Arg Val His 425 430 435	1349
tat ggc ttg tac aac cag tgc gac ccc tgc cct gga gag ttc ctc tgt Tyr Gly Leu Tyr Asn Gln Ser Asp Pro Cys Pro Gly Glu Phe Leu Cys 440 445 450 455	1397
tct gtg aat gga ctc tgt gtc cct gcc tgt gat ggg gtc aag gac tgc Ser Val Asn Gly Leu Cys Val Pro Ala Cys Asp Gly Val Lys Asp Cys 460 465 470	1445
ccc aac ggc ctg gat gag aga aac tgc gtt tgc aga gcc aca ttc cag Pro Asn Gly Leu Asp Glu Arg Asn Cys Val Cys Arg Ala Thr Phe Gln 475 480 485	1493
tgc aaa gag gac agc aca tgc atc tca ctg ccc aag gtc tgt gat ggg Cys Lys Glu Asp Ser Thr Cys Ile Ser Leu Pro Lys Val Cys Asp Gly 490 495 500	1541
cag cct gat tgt ctc aac ggc agc gac gaa gag cag tgc cag gaa ggg Gln Pro Asp Cys Leu Asn Gly Ser Asp Glu Glu Gln Cys Gln Glu Gly 505 510 515	1589
gtg cca tgt ggg aca ttc acc ttc cag tgt gag gac cgg agc tgc gtg Val Pro Cys Gly Thr Phe Thr Phe Gln Cys Glu Asp Arg Ser Cys Val 520 525 530 535	1637
aag aag ccc aac ccg cag tgt gat ggg cgg ccc gac tgc agg gac ggc Lys Lys Pro Asn Pro Gln Cys Asp Gly Arg Pro Asp Cys Arg Asp Gly 540 545 550	1685
tcg gat gag gag cac tgt gaa tgt ggc ctc cag ggc ccc tcc agc cgc Ser Asp Glu Glu His Cys Glu Cys Gly Leu Gln Gly Pro Ser Ser Arg 555 560 565	1733
att gtt ggt gga gct gtg tcc tcc gag ggt gag tgg cca tgg cag gcc Ile Val Gly Gly Ala Val Ser Ser Glu Gly Glu Trp Pro Trp Gln Ala 570 575 580	1781
agc ctc cag gtt cgg ggt cga cac atc tgt ggg ggg gcc ctc atc gct Ser Leu Gln Val Arg Gly Arg His Ile Cys Gly Gly Ala Leu Ile Ala 585 590 595	1829
gac cgc tgg gtg ata aca gct gcc cac tgc ttc cag gag gac agc atg	1877

Asp 600	Arg	Trp	Val	Ile	Thr 605	Ala	Ala	His	Cys	Phe 610	Gln	Glu	Asp	Ser	Met 615																																																						
gcc Ala	tcc Ser	acg Thr	gtg Val	ctg Leu 620	tgg Trp	acc Thr	gtg Val	ttc Phe	ctg Leu 625	ggc Gly	aag Lys	gtg Val	tgg Trp	cag Gln 630	aac Asn	1925																																																					
tcg Ser	cgc Arg	tgg Trp	cct Pro 635	gga Gly	gag Glu	gtg Val	tcc Ser	ttc Phe 640	aag Lys	gtg Val	agc Ser	cgc Arg	ctg Leu 645	ctc Leu	ctg Leu	1973																																																					
cac His	ccg Pro	tac Tyr 650	cac His	gaa Glu	gag Glu	gac Asp	agc Ser 655	cat His	gac Asp	tac Tyr	gac Asp	gtg Val 660	gcg Ala	ctg Leu	ctg Leu	2021																																																					
cag Gln	ctc Leu 665	gac Asp	cac His	ccg Pro	gtg Val 670	gtg Val	cgc Arg	tcg Ser	gcc Ala	gcc Ala	gtg Val 675	cgc Arg	ccc Pro	gtc Val	tgc Cys	2069																																																					
ctg Leu 680	ccc Pro	gcg Ala	cgc Arg	tcc Ser	cac His 685	ttc Phe	ttc Phe	gag Glu	ccc Pro	ggc Gly 690	ctg Leu	cac His	tgc Cys	tgg Trp	att Ile 695	2117																																																					
acg Thr	ggc Gly	tgg Trp	ggc Gly	gcc Ala 700	ttg Leu	cgc Arg	gag Glu	ggc Gly	ggc Gly 705	ccc Pro	atc Ile	agc Ser	aac Asn	gct Ala 710	ctg Leu	2165																																																					
cag Gln	aaa Lys	gtg Val	gat Asp 715	gtg Val	cag Gln	ttg Leu	atc Ile	cca Pro 720	cag Gln	gac Asp	ctg Leu	tgc Cys	agc Ser 725	gag Glu	gtc Val	2213																																																					
tat Tyr	cgc Arg	tac Tyr 730	cag Gln	gtg Val	acg Thr	cca Pro	cgc Arg 735	atg Met	ctg Leu	tgt Cys	gcc Ala	ggc Gly 740	tac Tyr	cgc Arg	aag Lys	2261																																																					
ggc Gly	aag Lys 745	aag Lys	gat Asp	gcc Ala	tgt Cys	cag Gln	ggt Gly	gac Asp	tca Ser	ggt Gly	ggt Gly 755	ccg Pro	ctg Leu	gtg Val	tgc Cys	2309																																																					
aag Lys 760	gca Ala	ctc Leu	agt Ser	ggc Gly	cgc Arg 765	tgg Trp	ttc Phe	ctg Leu	gcg Ala	ggg Gly 770	ctg Leu	gtc Val	agc Ser	tgg Trp	ggc Gly 775	2357																																																					
ctg Leu	ggc Gly	tgt Cys	ggc Gly	cgg Arg 780	cct Pro	aac Asn	tac Tyr	ttc Phe	ggc Gly 785	gtc Val	tac Tyr	acc Thr	cgc Arg	atc Ile 790	aca Thr	2405																																																					
ggt Gly	gtg Val	atc Ile	agc Ser	tgg Trp	atc Ile	cag Gln	caa Gln	gtg Val	gtg Val	acc Thr	tga *	ggaactgccc				2451																																																					
ccctgcaaag	caggGCCcac	ctcctggact	cagagagccc	agggcAactg	ccaagcaggg	2511	ggacaagtat	tctggCGGGG	ggtgggggag	agagcaggcc	ctgtggtggc	aggaggggca	2571	tcttgTTTTcg	tccctgatgt	ctgtCCagta	tggcaggagg	atgagaagtg	ccagcagttg	2631	ggggTCaaga	cgtccCTtGa	ggacCCcaggc	ccacacCCcag	ccctTTtTGcc	tccCAattct	2691	ctctcctccg	TccccTtccT	ccactgctgc	ctaAtGCaaG	gcagtggctc	agcagCaaga	2751	atgctggTtc	tacatCCCga	ggagtgtctg	aggtgcgcc	cactctgtac	agaggctgtt	2811	tgggcagcct	tgcctCcaga	gagcagattc	cagcttcCGga	agCCCCtggT	ctaacttggg	2871	atctgggaat	ggaaggTgct	cccataCGgag	gggacCCctca	gagCCctggga	gactgccagg	2931	tgggcctgct	gccactGTaa	gCCAAAaggT	ggggAAgTcc	tgaCTccagg	gtccttGccc	2991	cacctctgcc	tgccacctgg	gcctcAcag	cccAgacCCt	cactgggagg	tgagctcagc	3051

-16-

tgcccttttg aataaagctg cctgatgcaa aaaaaaaaaa aaaaaaaaaa aaa

3104

<210> 8  
 <211> 802  
 <212> PRT  
 <213> Homo Sapien

<400> 8  
 Met Pro Val Ala Glu Ala Pro Gln Val Ala Gly Gly Gln Gly Asp Gly  
 1 5 10 15  
 Gly Asp Gly Glu Glu Ala Glu Pro Glu Gly Met Phe Lys Ala Cys Glu  
 20 25 30  
 Asp Ser Lys Arg Lys Ala Arg Gly Tyr Leu Arg Leu Val Pro Leu Phe  
 35 40 45  
 Val Leu Leu Ala Leu Leu Val Leu Ala Ser Ala Gly Val Leu Leu Trp  
 50 55 60  
 Tyr Phe Leu Gly Tyr Lys Ala Glu Val Met Val Ser Gln Val Tyr Ser  
 65 70 75 80  
 Gly Ser Leu Arg Val Leu Asn Arg His Phe Ser Gln Asp Leu Thr Arg  
 85 90 95  
 Arg Glu Ser Ser Ala Phe Arg Ser Glu Thr Ala Lys Ala Gln Lys Met  
 100 105 110  
 Leu Lys Glu Leu Ile Thr Ser Thr Arg Leu Gly Thr Tyr Tyr Asn Ser  
 115 120 125  
 Ser Ser Val Tyr Ser Phe Gly Glu Gly Pro Leu Thr Cys Phe Phe Trp  
 130 135 140  
 Phe Ile Leu Gln Ile Pro Glu His Arg Arg Leu Met Leu Ser Pro Glu  
 145 150 155 160  
 Val Val Gln Ala Leu Leu Val Glu Glu Leu Ser Thr Val Asn Ser  
 165 170 175  
 Ser Ala Ala Val Pro Tyr Arg Ala Glu Tyr Glu Val Asp Pro Glu Gly  
 180 185 190  
 Leu Val Ile Leu Glu Ala Ser Val Lys Asp Ile Ala Ala Leu Asn Ser  
 195 200 205  
 Thr Leu Gly Cys Tyr Arg Tyr Ser Tyr Val Gly Gln Gly Gln Val Leu  
 210 215 220  
 Arg Leu Lys Gly Pro Asp His Leu Ala Ser Ser Cys Leu Trp His Leu  
 225 230 235 240  
 Gln Gly Pro Lys Asp Leu Met Leu Lys Leu Arg Leu Glu Trp Thr Leu  
 245 250 255  
 Ala Glu Cys Arg Asp Arg Leu Ala Met Tyr Asp Val Ala Gly Pro Leu  
 260 265 270  
 Glu Lys Arg Leu Ile Thr Ser Val Tyr Gly Cys Ser Arg Gln Glu Pro  
 275 280 285  
 Val Val Glu Val Leu Ala Ser Gly Ala Ile Met Ala Val Val Trp Lys  
 290 295 300  
 Lys Gly Leu His Ser Tyr Tyr Asp Pro Phe Val Leu Ser Val Gln Pro  
 305 310 315 320  
 Val Val Phe Gln Ala Cys Glu Val Asn Leu Thr Leu Asp Asn Arg Leu  
 325 330 335  
 Asp Ser Gln Gly Val Leu Ser Thr Pro Tyr Phe Pro Ser Tyr Tyr Ser  
 340 345 350  
 Pro Gln Thr His Cys Ser Trp His Leu Thr Val Pro Ser Leu Asp Tyr  
 355 360 365  
 Gly Leu Ala Leu Trp Phe Asp Ala Tyr Ala Leu Arg Arg Gln Lys Tyr  
 370 375 380  
 Asp Leu Pro Cys Thr Gln Gly Gln Trp Thr Ile Gln Asn Arg Arg Leu  
 385 390 395 400  
 Cys Gly Leu Arg Ile Leu Gln Pro Tyr Ala Glu Arg Ile Pro Val Val  
 405 410 415  
 Ala Thr Ala Gly Ile Thr Ile Asn Phe Thr Ser Gln Ile Ser Leu Thr

-17-

Gly	Pro	Gly	Val	Arg	Val	His	Tyr	Gly	Leu	Tyr	Asn	Gln	Ser	Asp	Pro
		435					440					445			
Cys	Pro	Gly	Glu	Phe	Leu	Cys	Ser	Val	Asn	Gly	Leu	Cys	Val	Pro	Ala
	450					455					460				
Cys	Asp	Gly	Val	Lys	Asp	Cys	Pro	Asn	Gly	Leu	Asp	Glu	Arg	Asn	Cys
465					470					475					480
Val	Cys	Arg	Ala	Thr	Phe	Gln	Cys	Lys	Glu	Asp	Ser	Thr	Cys	Ile	Ser
				485					490					495	
Leu	Pro	Lys	Val	Cys	Asp	Gly	Gln	Pro	Asp	Cys	Leu	Asn	Gly	Ser	Asp
			500					505					510		
Glu	Glu	Gln	Cys	Gln	Glu	Gly	Val	Pro	Cys	Gly	Thr	Phe	Thr	Phe	Gln
		515					520					525			
Cys	Glu	Asp	Arg	Ser	Cys	Val	Lys	Lys	Pro	Asn	Pro	Gln	Cys	Asp	Gly
	530					535				540					
Arg	Pro	Asp	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Glu	His	Cys	Glu	Cys	Gly
545					550					555					560
Leu	Gln	Gly	Pro	Ser	Ser	Arg	Ile	Val	Gly	Gly	Ala	Val	Ser	Ser	Glu
				565					570					575	
Gly	Glu	Trp	Pro	Trp	Gln	Ala	Ser	Leu	Gln	Val	Arg	Gly	Arg	His	Ile
			580					585					590		
Cys	Gly	Gly	Ala	Leu	Ile	Ala	Asp	Arg	Trp	Val	Ile	Thr	Ala	Ala	His
		595					600					605			
Cys	Phe	Gln	Glu	Asp	Ser	Met	Ala	Ser	Thr	Val	Leu	Trp	Thr	Val	Phe
	610					615					620				
Leu	Gly	Lys	Val	Trp	Gln	Asn	Ser	Arg	Trp	Pro	Gly	Glu	Val	Ser	Phe
625					630					635					640
Lys	Val	Ser	Arg	Leu	Leu	His	Pro	Tyr	His	Glu	Glu	Asp	Ser	His	
				645				650					655		
Asp	Tyr	Asp	Val	Ala	Leu	Leu	Gln	Leu	Asp	His	Pro	Val	Val	Arg	Ser
			660					665					670		
Ala	Ala	Val	Arg	Pro	Val	Cys	Leu	Pro	Ala	Arg	Ser	His	Phe	Phe	Glu
		675					680					685			
Pro	Gly	Leu	His	Cys	Trp	Ile	Thr	Gly	Trp	Gly	Ala	Leu	Arg	Glu	Gly
	690					695					700				
Gly	Pro	Ile	Ser	Asn	Ala	Leu	Gln	Lys	Val	Asp	Val	Gln	Leu	Ile	Pro
705					710					715					720
Gln	Asp	Leu	Cys	Ser	Glu	Val	Tyr	Arg	Tyr	Gln	Val	Thr	Pro	Arg	Met
				725					730					735	
Leu	Cys	Ala	Gly	Tyr	Arg	Lys	Gly	Lys	Asp	Ala	Cys	Gln	Gly	Asp	
			740					745					750		
Ser	Gly	Gly	Pro	Leu	Val	Cys	Lys	Ala	Leu	Ser	Gly	Arg	Trp	Phe	Leu
	755						760					765			
Ala	Gly	Leu	Val	Ser	Trp	Gly	Leu	Gly	Cys	Gly	Arg	Pro	Asn	Tyr	Phe
	770					775					780				
Gly	Val	Tyr	Thr	Arg	Ile	Thr	Gly	Val	Ile	Ser	Trp	Ile	Gln	Gln	Val
785					790					795					800
Val	Thr														

&lt;210&gt; 9

&lt;211&gt; 2672

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (33)...(2009)

&lt;223&gt; cDNA encoding: MTSP4-S (short form) splice variant

&lt;400&gt; 9

-18-

tcatacggcca gaggggtgatc agtgagcaga ag atg ccc gtg gcc gag gcc ccc	53
Met Pro Val Ala Glu Ala Pro	
1 5	
cag gtg gct ggc ggg cag ggg gac gga ggt gat ggc gag gaa gcg gag	101
Gln Val Ala Gly Gly Gln Gly Asp Gly Gly Asp Gly Glu Glu Ala Glu	
10 15 20	
ccg gag ggg atg ttc aag gcc tgt gag gac tcc aag aga aaa gcc cgg	149
Pro Glu Gly Met Phe Lys Ala Cys Glu Asp Ser Lys Arg Lys Ala Arg	
25 30 35	
ggc tac ctc cgc ctg gtg ccc ctg ttt gtg ctg ctg gcc ctg ctc gtg	197
Gly Tyr Leu Arg Leu Val Pro Leu Phe Val Leu Leu Ala Leu Leu Val	
40 45 50 55	
ctg gct tcg gcg ggg gtg cta ctc tgg tat ttc cta ggg tac aag gcg	245
Leu Ala Ser Ala Gly Val Leu Leu Trp Tyr Phe Leu Gly Tyr Lys Ala	
60 65 70	
gag gtg atg gtc agc cag gtg tac tca ggc agt ctg cgt gta ctc aat	293
Glu Val Met Val Ser Gln Val Tyr Ser Gly Ser Leu Arg Val Leu Asn	
75 80 85	
cgc cac ttc tcc cag gat ctt acc cgc cgg gaa tct agt gcc ttc cgc	341
Arg His Phe Ser Gln Asp Leu Thr Arg Arg Glu Ser Ser Ala Phe Arg	
90 95 100	
agt gaa acc gcc aaa gcc cag aag atg ctc aag gag ctc atc acc agc	389
Ser Glu Thr Ala Lys Ala Gln Lys Met Leu Lys Glu Leu Ile Thr Ser	
105 110 115	
acc cgc ctg gga act tac tac aac tcc agc tcc gtc tat tcc ttt ggg	437
Thr Arg Leu Gly Thr Tyr Tyr Asn Ser Ser Ser Val Tyr Ser Phe Gly	
120 125 130 135	
gtg tac ggc tgc agc cgc cag gag ccc gtg gtg gag gtt ctg gcg tcg	485
Val Tyr Gly Cys Ser Arg Gln Glu Pro Val Val Glu Val Leu Ala Ser	
140 145 150	
ggg gcc atc atg gcg gtc gtc tgg aag aag ggc ctg cac agc tac tac	533
Gly Ala Ile Met Ala Val Val Trp Lys Lys Gly Leu His Ser Tyr Tyr	
155 160 165	
gac ccc ttc gtg ctc tcc gtg cag ccg gtg gtc ttc cag gcc tgt gaa	581
Asp Pro Phe Val Leu Ser Val Gln Pro Val Val Phe Gln Ala Cys Glu	
170 175 180	
gtg aac ctg acg ctg gac aac agg ctc gac tcc cag ggc gtc ctc agc	629
Val Asn Leu Thr Leu Asp Asn Arg Leu Asp Ser Gln Gly Val Leu Ser	
185 190 195	
acc ccg tac ttc ccc agc tac tac tcg ccc caa acc cac tgc tcc tgg	677
Thr Pro Tyr Phe Pro Ser Tyr Tyr Ser Pro Gln Thr His Cys Ser Trp	
200 205 210 215	
cac ctc acg gtg ccc tct ctg gac tac ggc ttg gcc ctc tgg ttt gat	725
His Leu Thr Val Pro Ser Leu Asp Tyr Gly Leu Ala Leu Trp Phe Asp	
220 225 230	
gcc tat gca ctg agg agg cag aag tat gat ttg ccg tgc acc cag ggc	773

-19-

Ala	Tyr	Ala	Leu	Arg	Arg	Gln	Lys	Tyr	Asp	Leu	Pro	Cys	Thr	Gln	Gly	
			235					240					245			
cag	tgg	acg	atc	cag	aac	agg	agg	ctg	tgt	ggc	ttg	cgc	atc	ctg	cag	821
Gln	Trp	Thr	Ile	Gln	Asn	Arg	Arg	Leu	Cys	Gly	Leu	Arg	Ile	Leu	Gln	
		250					255					260				
ccc	tac	gcc	gag	agg	atc	ccc	gtg	gtg	gcc	acg	gcc	ggg	atc	acc	atc	869
Pro	Tyr	Ala	Glu	Arg	Ile	Pro	Val	Val	Ala	Thr	Ala	Gly	Ile	Thr	Ile	
	265					270					275					
aac	ttc	acc	tcc	cag	atc	tcc	ctc	acc	ggg	ccc	ggg	gtg	cgg	gtg	cac	917
Asn	Phe	Thr	Ser	Gln	Ile	Ser	Leu	Thr	Gly	Pro	Gly	Val	Arg	Val	His	
280					285				290						295	
tat	ggc	ttg	tac	aac	cag	tcg	gac	ccc	tgc	cct	gga	gag	ttc	ctc	tgt	965
Tyr	Gly	Leu	Tyr	Asn	Gln	Ser	Asp	Pro	Cys	Pro	Gly	Glu	Phe	Leu	Cys	
				300					305					310		
tct	gtg	aat	gga	ctc	tgt	gtc	cct	gcc	tgt	gat	ggg	gtc	aag	gac	tgc	1013
Ser	Val	Asn	Gly	Leu	Cys	Val	Pro	Ala	Cys	Asp	Gly	Val	Lys	Asp	Cys	
			315					320					325			
ccc	aac	ggc	ctg	gat	gag	aga	aac	tgc	gtt	tgc	aga	gcc	aca	ttc	cag	1061
Pro	Asn	Gly	Leu	Asp	Glu	Arg	Asn	Cys	Val	Cys	Arg	Ala	Thr	Phe	Gln	
		330					335					340				
tgc	aaa	gag	gac	agc	aca	tgc	atc	tca	ctg	ccc	aag	gtc	tgt	gat	ggg	1109
Cys	Lys	Glu	Asp	Ser	Thr	Cys	Ile	Ser	Leu	Pro	Lys	Val	Cys	Asp	Gly	
	345					350					355					
cag	cct	gat	tgt	ctc	aac	ggc	agc	gac	gaa	gag	cag	tgc	cag	gaa	ggg	1157
Gln	Pro	Asp	Cys	Leu	Asn	Gly	Ser	Asp	Glu	Glu	Gln	Cys	Gln	Glu	Gly	
360					365				370					375		
gtg	cca	tgt	ggg	aca	ttc	acc	ttc	cag	tgt	gag	gac	cgg	agc	tgc	gtg	1205
Val	Pro	Cys	Gly	Thr	Phe	Thr	Phe	Gln	Cys	Glu	Asp	Arg	Ser	Cys	Val	
				380				385						390		
aag	aag	ccc	aac	ccg	cag	tgt	gat	ggg	cgg	ccc	gac	tgc	agg	gac	ggc	1253
Lys	Lys	Pro	Asn	Pro	Gln	Cys	Asp	Gly	Arg	Pro	Asp	Cys	Arg	Asp	Gly	
			395				400					405				
tcg	gat	gag	gag	cac	tgt	gaa	tgt	ggc	ctc	cag	ggc	ccc	tcc	agc	cgc	1301
Ser	Asp	Glu	Glu	His	Cys	Glu	Cys	Gly	Leu	Gln	Gly	Pro	Ser	Ser	Arg	
		410				415						420				
att	gtt	ggg	gga	gct	gtg	tcc	tcc	gag	ggg	gag	tgg	cca	tgg	cag	gcc	1349
Ile	Val	Gly	Gly	Ala	Val	Ser	Ser	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Ala	
		425				430					435					
agc	ctc	cag	gtt	cgg	ggg	cga	cac	atc	tgt	ggg	ggg	gcc	ctc	atc	gct	1397
Ser	Leu	Gln	Val	Arg	Gly	Arg	His	Ile	Cys	Gly	Gly	Ala	Leu	Ile	Ala	
440					445				450						455	
gac	cgc	tgg	gtg	ata	aca	gct	gcc	cac	tgc	ttc	cag	gag	gac	agc	atg	1445
Asp	Arg	Trp	Val	Ile	Thr	Ala	Ala	His	Cys	Phe	Gln	Glu	Asp	Ser	Met	
				460				465						470		
gcc	tcc	acg	gtg	ctg	tgg	acc	gtg	ttc	ctg	ggc	aag	gtg	tgg	cag	aac	1493
Ala	Ser	Thr	Val	Leu	Trp	Thr	Val	Phe	Leu	Gly	Lys	Val	Trp	Gln	Asn	

-20-

475	480	485	
tgc cgc tgg cct gga gag gtg tcc ttc aag gtg agc cgc ctg ctc ctg			1541
Ser Arg Trp Pro Gly Glu Val Ser Phe Lys Val Ser Arg Leu Leu Leu			
490	495	500	
cac ccg tac cac gaa gag gac agc cat gac tac gac gtg gcg ctg ctg			1589
His Pro Tyr His Glu Glu Asp Ser His Asp Tyr Asp Val Ala Leu Leu			
505	510	515	
cag ctc gac cac ccg gtg gtg cgc tgc gcc gcc gtg cgc ccc gtc tgc			1637
Gln Leu Asp His Pro Val Val Arg Ser Ala Ala Val Arg Pro Val Cys			
520	525	530	535
ctg ccc gcg cgc tcc cac ttc ttc gag ccc gcc ctg cac tgc tgg att			1685
Leu Pro Ala Arg Ser His Phe Phe Glu Pro Gly Leu His Cys Trp Ile			
	540	545	550
acg gcc tgg gcc gcc ttg cgc gag gcc gcc ccc atc agc aac gct ctg			1733
Thr Gly Trp Gly Ala Leu Arg Glu Gly Gly Pro Ile Ser Asn Ala Leu			
	555	560	565
cag aaa gtg gat gtg cag ttg atc cca cag gac ctg tgc agc gag gtc			1781
Gln Lys Val Asp Val Gln Leu Ile Pro Gln Asp Leu Cys Ser Glu Val			
	570	575	580
tat cgc tac cag gtg acg cca cgc atg ctg tgt gcc gcc tac cgc aag			1829
Tyr Arg Tyr Gln Val Thr Pro Arg Met Leu Cys Ala Gly Tyr Arg Lys			
	585	590	595
ggc aag aag gat gcc tgt cag ggt gac tca ggt ggt ccg ctg gtg tgc			1877
Gly Lys Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys			
600	605	610	615
aag gca ctc agt gcc cgc tgg ttc ctg gcg ggg ctg gtc agc tgg gcc			1925
Lys Ala Leu Ser Gly Arg Trp Phe Leu Ala Gly Leu Val Ser Trp Gly			
	620	625	630
ctg gcc tgt gcc cgg cct aac tac ttc gcc gtc tac acc cgc atc aca			1973
Leu Gly Cys Gly Arg Pro Asn Tyr Phe Gly Val Tyr Thr Arg Ile Thr			
	635	640	645
ggt gtg atc agc tgg atc cag caa gtg gtg acc tga ggaactgccc			2019
Gly Val Ile Ser Trp Ile Gln Gln Val Val Thr *			
650	655		
ccctgcaaag cagggccac ctctggact cagagagccc agggcaactg ccaagcaggg			2079
ggacaagtat tctggcggg ggtgggggag agagcaggcc ctgtggtggc aggaggggca			2139
tcttgtttcg tccctgatgt ctgtccagta tggcaggagg atgagaagtg ccagcagttg			2199
ggggcacaaga cgtcccttga ggaccagggc ccacacccag cccttttgcc tcccaattct			2259
ctctcctccg tcccttctt ccaactgctgc ctaatgcaag gcagtggctc agcagcaaga			2319
atgctgggttc tacatcccga ggagtgtctg aggtgcgccc cactctgtac agaggctgtt			2379
tgggcagcct tgcctccaga gagcagattc cagcttcgga agcccctggg ctaacttggg			2439
atctgggaat ggaaggtgct cccatcggag gggaccctca gagccctgga gactgccagg			2499
tgggcctgct gccactgtaa gccaaaaggt ggggaagtcc tgactccagg gtccttgccc			2559
cacccctgcc tgccacctgg gccctcacag ccagaccct cactgggagg tgagctcagc			2619
tgccctttgg aataaagctg cctgatgcaa aaaaaaaaaa aaaaaaaaaa aaa			2672
<210> 10			
<211> 658			
<212> PRT			

-21-

&lt;213&gt; Homo Sapien

&lt;400&gt; 10

Met	Pro	Val	Ala	Glu	Ala	Pro	Gln	Val	Ala	Gly	Gly	Gln	Gly	Asp	Gly
1				5					10					15	
Gly	Asp	Gly	Glu	Glu	Ala	Glu	Pro	Glu	Gly	Met	Phe	Lys	Ala	Cys	Glu
			20					25					30		
Asp	Ser	Lys	Arg	Lys	Ala	Arg	Gly	Tyr	Leu	Arg	Leu	Val	Pro	Leu	Phe
		35					40					45			
Val	Leu	Leu	Ala	Leu	Leu	Val	Leu	Ala	Ser	Ala	Gly	Val	Leu	Leu	Trp
	50					55					60				
Tyr	Phe	Leu	Gly	Tyr	Lys	Ala	Glu	Val	Met	Val	Ser	Gln	Val	Tyr	Ser
65					70					75					80
Gly	Ser	Leu	Arg	Val	Leu	Asn	Arg	His	Phe	Ser	Gln	Asp	Leu	Thr	Arg
				85					90					95	
Arg	Glu	Ser	Ser	Ala	Phe	Arg	Ser	Glu	Thr	Ala	Lys	Ala	Gln	Lys	Met
			100					105					110		
Leu	Lys	Glu	Leu	Ile	Thr	Ser	Thr	Arg	Leu	Gly	Thr	Tyr	Tyr	Asn	Ser
		115					120					125			
Ser	Ser	Val	Tyr	Ser	Phe	Gly	Val	Tyr	Gly	Cys	Ser	Arg	Gln	Glu	Pro
	130					135					140				
Val	Val	Glu	Val	Leu	Ala	Ser	Gly	Ala	Ile	Met	Ala	Val	Val	Trp	Lys
145					150					155					160
Lys	Gly	Leu	His	Ser	Tyr	Tyr	Asp	Pro	Phe	Val	Leu	Ser	Val	Gln	Pro
				165					170					175	
Val	Val	Phe	Gln	Ala	Cys	Glu	Val	Asn	Leu	Thr	Leu	Asp	Asn	Arg	Leu
			180					185					190		
Asp	Ser	Gln	Gly	Val	Leu	Ser	Thr	Pro	Tyr	Phe	Pro	Ser	Tyr	Tyr	Ser
		195					200					205			
Pro	Gln	Thr	His	Cys	Ser	Trp	His	Leu	Thr	Val	Pro	Ser	Leu	Asp	Tyr
	210					215					220				
Gly	Leu	Ala	Leu	Trp	Phe	Asp	Ala	Tyr	Ala	Leu	Arg	Arg	Gln	Lys	Tyr
225					230					235					240
Asp	Leu	Pro	Cys	Thr	Gln	Gly	Gln	Trp	Thr	Ile	Gln	Asn	Arg	Arg	Leu
				245					250					255	
Cys	Gly	Leu	Arg	Ile	Leu	Gln	Pro	Tyr	Ala	Glu	Arg	Ile	Pro	Val	Val
			260					265					270		
Ala	Thr	Ala	Gly	Ile	Thr	Ile	Asn	Phe	Thr	Ser	Gln	Ile	Ser	Leu	Thr
		275					280					285			
Gly	Pro	Gly	Val	Arg	Val	His	Tyr	Gly	Leu	Tyr	Asn	Gln	Ser	Asp	Pro
	290					295					300				
Cys	Pro	Gly	Glu	Phe	Leu	Cys	Ser	Val	Asn	Gly	Leu	Cys	Val	Pro	Ala
305					310					315					320
Cys	Asp	Gly	Val	Lys	Asp	Cys	Pro	Asn	Gly	Leu	Asp	Glu	Arg	Asn	Cys
				325					330					335	
Val	Cys	Arg	Ala	Thr	Phe	Gln	Cys	Lys	Glu	Asp	Ser	Thr	Cys	Ile	Ser
			340					345					350		
Leu	Pro	Lys	Val	Cys	Asp	Gly	Gln	Pro	Asp	Cys	Leu	Asn	Gly	Ser	Asp
		355					360					365			
Glu	Glu	Gln	Cys	Gln	Glu	Gly	Val	Pro	Cys	Gly	Thr	Phe	Thr	Phe	Gln
	370					375					380				
Cys	Glu	Asp	Arg	Ser	Cys	Val	Lys	Lys	Pro	Asn	Pro	Gln	Cys	Asp	Gly
385					390					395					400
Arg	Pro	Asp	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Glu	His	Cys	Glu	Cys	Gly
				405					410					415	
Leu	Gln	Gly	Pro	Ser	Ser	Arg	Ile	Val	Gly	Gly	Ala	Val	Ser	Ser	Glu
			420					425					430		
Gly	Glu	Trp	Pro	Trp	Gln	Ala	Ser	Leu	Gln	Val	Arg	Gly	Arg	His	Ile
		435					440					445			
Cys	Gly	Gly	Ala	Leu	Ile	Ala	Asp	Arg	Trp	Val	Ile	Thr	Ala	Ala	His
	450					455					460				



-22-

Cys Phe Gln Glu Asp Ser Met Ala Ser Thr Val Leu Trp Thr Val Phe  
 465 470 475 480  
 Leu Gly Lys Val Trp Gln Asn Ser Arg Trp Pro Gly Glu Val Ser Phe  
 485 490 495  
 Lys Val Ser Arg Leu Leu Leu His Pro Tyr His Glu Glu Asp Ser His  
 500 505 510  
 Asp Tyr Asp Val Ala Leu Leu Gln Leu Asp His Pro Val Val Arg Ser  
 515 520 525  
 Ala Ala Val Arg Pro Val Cys Leu Pro Ala Arg Ser His Phe Phe Glu  
 530 535 540  
 Pro Gly Leu His Cys Trp Ile Thr Gly Trp Gly Ala Leu Arg Glu Gly  
 545 550 555 560  
 Gly Pro Ile Ser Asn Ala Leu Gln Lys Val Asp Val Gln Leu Ile Pro  
 565 570 575  
 Gln Asp Leu Cys Ser Glu Val Tyr Arg Tyr Gln Val Thr Pro Arg Met  
 580 585 590  
 Leu Cys Ala Gly Tyr Arg Lys Gly Lys Lys Asp Ala Cys Gln Gly Asp  
 595 600 605  
 Ser Gly Gly Pro Leu Val Cys Lys Ala Leu Ser Gly Arg Trp Phe Leu  
 610 615 620  
 Ala Gly Leu Val Ser Trp Gly Leu Gly Cys Gly Arg Pro Asn Tyr Phe  
 625 630 635 640  
 Gly Val Tyr Thr Arg Ile Thr Gly Val Ile Ser Trp Ile Gln Gln Val  
 645 650 655  
 Val Thr

<210> 11  
 <211> 1656  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <221> CDS  
 <222> (268)...(1629)  
 <223> Nucleic acid encoding a transmembrane serine  
 protease (MTSP-6) protein

<400> 11  
 cgcccgggca ggtcagtaac actgtggcct actatctctt ccgtgggtgcc atctacattt 60  
 ttgggactcg ggaattatga ctgttttttg ttaatcgata ctgaatgcgc tttgtgtgga 120  
 ctgtcgaatt tcaaagattt accgtatgac caagatgcac ctgatgctac aagtataaat 180  
 aggggaacaa atgctttctg ttcttcctcg gctaaggagg tagaggtgga ggcggagccg 240  
 gatgtcagag gtcctgaaat agtcacc atg ggg gaa aat gat ccg cct gct gtt 294  
 Met Gly Glu Asn Asp Pro Pro Ala Val  
 1 5  
  
 gaa gcc ccc ttc tca ttc cga tgc ctt ttt ggc ctt gat gat ttg aaa 342  
 Glu Ala Pro Phe Ser Phe Arg Ser Leu Phe Gly Leu Asp Asp Leu Lys 25  
 10 15 20  
  
 ata agt cct gtt gca cca gat gca gat gct gtt gct gca cag atc ctg 390  
 Ile Ser Pro Val Ala Pro Asp Ala Asp Ala Val Ala Ala Gln Ile Leu 40  
 30 35  
  
 tca ctg ctg cca ttg aag ttt ttt cca atc atc gtc att ggg atc att 438  
 Ser Leu Leu Pro Leu Lys Phe Phe Pro Ile Ile Val Ile Gly Ile Ile 55  
 45 50  
  
 gca ttg ata tta gca ctg gcc att ggt ctg ggc atc cac ttc gac tgc 486  
 Ala Leu Ile Leu Ala Leu Ala Ile Gly Leu Gly Ile His Phe Asp Cys

-23-

60	65	70	
tca ggg aag tac aga tgt cgc tca tcc ttt aag tgt atc gag ctg ata			534
Ser Gly Lys Tyr Arg Cys Arg Ser Ser Phe Lys Cys Ile Glu Leu Ile			
75	80	85	
gct cga tgt gac gga gtc tcg gat tgc aaa gac ggg gag gac gag tac			582
Ala Arg Cys Asp Gly Val Ser Asp Cys Lys Asp Gly Glu Asp Glu Tyr			
90	95	100	105
cgc tgt gtc cgg gtg ggt ggt cag aat gcc gtg ctc cag gtg ttc aca			630
Arg Cys Val Arg Val Gly Gly Gln Asn Ala Val Leu Gln Val Phe Thr			
110	115	120	
gct gct tcg tgg aag acc atg tgc tcc gat gac tgg aag ggt cac tac			678
Ala Ala Ser Trp Lys Thr Met Cys Ser Asp Asp Trp Lys Gly His Tyr			
125	130	135	
gca aat gtt gcc tgt gcc caa ctg ggt ttc cca agc tat gta agt tca			726
Ala Asn Val Ala Cys Ala Gln Leu Gly Phe Pro Ser Tyr Val Ser Ser			
140	145	150	
gat aac ctc aga gtg agc tcg cta gag ggg cag ttc cgg gag gag ttt			774
Asp Asn Leu Arg Val Ser Ser Leu Glu Gly Gln Phe Arg Glu Glu Phe			
155	160	165	
gtg tcc atc gat cac ctc ttg cca gat gac aag gtg act gca tta cac			822
Val Ser Ile Asp His Leu Leu Pro Asp Asp Lys Val Thr Ala Leu His			
170	175	180	185
cac tca gta tat gtg agg gag gga tgt gcc tct ggc cac gtg gtt acc			870
His Ser Val Tyr Val Arg Glu Gly Cys Ala Ser Gly His Val Val Thr			
190	195	200	
ttg cag tgc aca gcc tgt ggt cat aga agg ggc tac agc tca cgc atc			918
Leu Gln Cys Thr Ala Cys Gly His Arg Arg Gly Tyr Ser Ser Arg Ile			
205	210	215	
gtg ggt gga aac atg tcc ttg ctc tcg cag tgg ccc tgg cag gcc agc			966
Val Gly Gly Asn Met Ser Leu Ser Gln Trp Pro Trp Gln Ala Ser			
220	225	230	
ctt cag ttc cag ggc tac cac ctg tgc ggg ggc tct gtc atc acg ccc			1014
Leu Gln Phe Gln Gly Tyr His Leu Cys Gly Gly Ser Val Ile Thr Pro			
235	240	245	
ctg tgg atc atc act gct gca cac tgt gtt tat gac ttg tac ctc ccc			1062
Leu Trp Ile Ile Thr Ala Ala His Cys Val Tyr Asp Leu Tyr Leu Pro			
250	255	260	265
aag tca tgg acc atc cag gtg ggt cta gtt tcc ctg ttg gac aat cca			1110
Lys Ser Trp Thr Ile Gln Val Gly Leu Val Ser Leu Leu Asp Asn Pro			
270	275	280	
gcc cca tcc cac ttg gtg gag aag att gtc tac cac agc aag tac aag			1158
Ala Pro Ser His Leu Val Glu Lys Ile Val Tyr His Ser Lys Tyr Lys			
285	290	295	
cca aag agg ctg ggc aat gac atc gcc ctt atg aag ctg gcc ggg cca			1206
Pro Lys Arg Leu Gly Asn Asp Ile Ala Leu Met Lys Leu Ala Gly Pro			
300	305	310	

-24-

ctc acg ttc aat gaa atg atc cag cct gtg tgc ctg ccc aac tct gaa Leu Thr Phe Asn Glu Met Ile Gln Pro Val Cys Leu Pro Asn Ser Glu 315 320 325	1254
gag aac ttc ccc gat gga aaa gtg tgc tgg acg tca gga tgg ggg gcc Glu Asn Phe Pro Asp Gly Lys Val Cys Trp Thr Ser Gly Trp Gly Ala 330 335 340 345	1302
aca gag gat gga ggt gac gcc tcc cct gtc ctg aac cac gcg gcc gtc Thr Glu Asp Gly Asp Ala Ser Pro Val Leu Asn His Ala Ala Val 350 355 360	1350
cct ttg att tcc aac aag atc tgc aac cac agg gac gtg tac ggt ggc Pro Leu Ile Ser Asn Lys Ile Cys Asn His Arg Asp Val Tyr Gly Gly 365 370 375	1398
atc atc tcc ccc tcc atg ctc tgc gcg ggc tac ctg acg ggt ggc gtg Ile Ile Ser Pro Ser Met Leu Cys Ala Gly Tyr Leu Thr Gly Gly Val 380 385 390	1446
gac agc tgc cag ggg gac agc ggg ggg ccc ctg gtg tgt caa gag agg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Glu Arg 395 400 405	1494
agg ctg tgg aag tta gtg gga gcg acc agc ttt ggc atc ggc tgc gca Arg Leu Trp Lys Leu Val Gly Ala Thr Ser Phe Gly Ile Gly Cys Ala 410 415 420 425	1542
gag gtg aac aag cct ggg gtg tac acc cgt gtc acc tcc ttc ctg gac Glu Val Asn Lys Pro Gly Val Tyr Thr Arg Val Thr Ser Phe Leu Asp 430 435 440	1590
tgg atc cac gag cag atg gag aga gac cta aaa acc tga agaggaaggg Trp Ile His Glu Gln Met Glu Arg Asp Leu Lys Thr * 445 450	1639
gataagtagc cacctga	1656
<210> 12	
<211> 453	
<212> PRT	
<213> Homo Sapien	
<400> 12	
Met Gly Glu Asn Asp Pro Pro Ala Val Glu Ala Pro Phe Ser Phe Arg 1 5 10 15	
Ser Leu Phe Gly Leu Asp Asp Leu Lys Ile Ser Pro Val Ala Pro Asp 20 25 30	
Ala Asp Ala Val Ala Ala Gln Ile Leu Ser Leu Leu Pro Leu Lys Phe 35 40 45	
Phe Pro Ile Ile Val Ile Gly Ile Ile Ala Leu Ile Leu Ala Leu Ala 50 55 60	
Ile Gly Leu Gly Ile His Phe Asp Cys Ser Gly Lys Tyr Arg Cys Arg 65 70 75 80	
Ser Ser Phe Lys Cys Ile Glu Leu Ile Ala Arg Cys Asp Gly Val Ser 85 90 95	
Asp Cys Lys Asp Gly Glu Asp Glu Tyr Arg Cys Val Arg Val Gly Gly 100 105 110	
Gln Asn Ala Val Leu Gln Val Phe Thr Ala Ala Ser Trp Lys Thr Met 115 120 125	

-25-

Cys Ser Asp Asp Trp Lys Gly His Tyr Ala Asn Val Ala Cys Ala Gln  
 130 135 140  
 Leu Gly Phe Pro Ser Tyr Val Ser Ser Asp Asn Leu Arg Val Ser Ser  
 145 150 155 160  
 Leu Glu Gly Gln Phe Arg Glu Glu Phe Val Ser Ile Asp His Leu Leu  
 165 170 175  
 Pro Asp Asp Lys Val Thr Ala Leu His His Ser Val Tyr Val Arg Glu  
 180 185 190  
 Gly Cys Ala Ser Gly His Val Val Thr Leu Gln Cys Thr Ala Cys Gly  
 195 200 205  
 His Arg Arg Gly Tyr Ser Ser Arg Ile Val Gly Gly Asn Met Ser Leu  
 210 215 220  
 Leu Ser Gln Trp Pro Trp Gln Ala Ser Leu Gln Phe Gln Gly Tyr His  
 225 230 235 240  
 Leu Cys Gly Gly Ser Val Ile Thr Pro Leu Trp Ile Ile Thr Ala Ala  
 245 250 255  
 His Cys Val Tyr Asp Leu Tyr Leu Pro Lys Ser Trp Thr Ile Gln Val  
 260 265 270  
 Gly Leu Val Ser Leu Leu Asp Asn Pro Ala Pro Ser His Leu Val Glu  
 275 280 285  
 Lys Ile Val Tyr His Ser Lys Tyr Lys Pro Lys Arg Leu Gly Asn Asp  
 290 295 300  
 Ile Ala Leu Met Lys Leu Ala Gly Pro Leu Thr Phe Asn Glu Met Ile  
 305 310 315 320  
 Gln Pro Val Cys Leu Pro Asn Ser Glu Glu Asn Phe Pro Asp Gly Lys  
 325 330 335  
 Val Cys Trp Thr Ser Gly Trp Gly Ala Thr Glu Asp Gly Gly Asp Ala  
 340 345 350  
 Ser Pro Val Leu Asn His Ala Ala Val Pro Leu Ile Ser Asn Lys Ile  
 355 360 365  
 Cys Asn His Arg Asp Val Tyr Gly Gly Ile Ile Ser Pro Ser Met Leu  
 370 375 380  
 Cys Ala Gly Tyr Leu Thr Gly Gly Val Asp Ser Cys Gln Gly Asp Ser  
 385 390 395 400  
 Gly Gly Pro Leu Val Cys Gln Glu Arg Arg Leu Trp Lys Leu Val Gly  
 405 410 415  
 Ala Thr Ser Phe Gly Ile Gly Cys Ala Glu Val Asn Lys Pro Gly Val  
 420 425 430  
 Tyr Thr Arg Val Thr Ser Phe Leu Asp Trp Ile His Glu Gln Met Glu  
 435 440 445  
 Arg Asp Leu Lys Thr  
 450

<210> 13  
 <211> 2100  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> CDS  
 <222> (45)...(1361)  
 <223> Nucleic acid encoding MTSP7

<400> 13  
 agatcagatg gcgactgaat agaagctgcc ccagtcctgg gttc atg atg tac aca 56  
 Met Met Tyr Thr  
 1

cct gtt gaa ttt tca gaa gct gaa ttc tca cga gct gaa tat caa aga 104  
 Pro Val Glu Phe Ser Glu Ala Glu Phe Ser Arg Ala Glu Tyr Gln Arg  
 5 10 15 20

-26-

aag cag caa ttt tgg gac tca gta cgg cta gct ctt ttc aca tta gca Lys Gln Gln Phe Trp Asp Ser Val Arg Leu Ala Leu Phe Thr Leu Ala 25 30 35	152
att gta gca atc ata gga att gca att ggt att gtt act cat ttt gtt Ile Val Ala Ile Ile Gly Ile Ala Ile Gly Ile Val Thr His Phe Val 40 45 50	200
gtt gag gat gat aag tct ttc tat tac ctt gcc tct ttt aaa gtc aca Val Glu Asp Asp Lys Ser Phe Tyr Tyr Leu Ala Ser Phe Lys Val Thr 55 60 65	248
aat atc aaa tat aaa gaa aat tat ggc ata aga tct tca aga gag ttt Asn Ile Lys Tyr Lys Glu Asn Tyr Gly Ile Arg Ser Ser Arg Glu Phe 70 75 80	296
ata gaa agg agt cat cag att gaa aga atg atg tct agg ata ttt cga Ile Glu Arg Ser His Gln Ile Glu Arg Met Met Ser Arg Ile Phe Arg 85 90 95 100	344
cat tct tct gta ggc ggt cga ttt atc aaa tct cat gtt atc aaa tta His Ser Ser Val Gly Gly Arg Phe Ile Lys Ser His Val Ile Lys Leu 105 110 115	392
agt cca gat gaa caa ggt gtg gat att ctt ata gtg ctc ata ttt cga Ser Pro Asp Glu Gln Gly Val Asp Ile Leu Ile Val Leu Ile Phe Arg 120 125 130	440
tac cca tct act gat agt gct gaa caa atc aag aaa aaa att gaa aag Tyr Pro Ser Thr Asp Ser Ala Glu Gln Ile Lys Lys Lys Ile Glu Lys 135 140 145	488
gct tta tat caa agt ttg aag acc aaa caa ttg tct ttg acc ata aac Ala Leu Tyr Gln Ser Leu Lys Thr Lys Gln Leu Ser Leu Thr Ile Asn 150 155 160	536
aaa cca tca ttt aga ctc aca cct att gac agc aaa aag atg agg aat Lys Pro Ser Phe Arg Leu Thr Pro Ile Asp Ser Lys Lys Met Arg Asn 165 170 175 180	584
ctt ctc aac agt cgc tgt gga ata agg atg aca tct tca aac atg cca Leu Leu Asn Ser Arg Cys Gly Ile Arg Met Thr Ser Ser Asn Met Pro 185 190 195	632
tta cca gca tcc tct tct act caa aga att gtc caa gga agg gaa aca Leu Pro Ala Ser Ser Ser Thr Gln Arg Ile Val Gln Gly Arg Glu Thr 200 205 210	680
gct atg gaa ggg gaa tgg cca tgg cag gcc agc ctc cag ctc ata ggg Ala Met Glu Gly Glu Trp Pro Trp Gln Ala Ser Leu Gln Leu Ile Gly 215 220 225	728
tca ggc cat cag tgt gga gcc agc ctc atc agt aac aca tgg ctg ctc Ser Gly His Gln Cys Gly Ala Ser Leu Ile Ser Asn Thr Trp Leu Leu 230 235 240	776
aca gca gct cac tgc ttt tgg aaa aat aaa gac cca act caa tgg att Thr Ala Ala His Cys Phe Trp Lys Asn Lys Asp Pro Thr Gln Trp Ile 245 250 255 260	824

-27-

gct act ttt ggt gca act ata aca cca ccc gca gtg aaa cga aat gtg Ala Thr Phe Gly Ala Thr Ile Thr Pro Pro Ala Val Lys Arg Asn Val	872
265 270 275	
agg aaa att att ctt cat gag aat tac cat aga gaa aca aat gaa aat Arg Lys Ile Ile Leu His Glu Asn Tyr His Arg Glu Thr Asn Glu Asn	920
280 285 290	
gac att gct ttg gtt cag ctc tct act gga gtt gag ttt tca aat ata Asp Ile Ala Leu Val Gln Leu Ser Thr Gly Val Glu Phe Ser Asn Ile	968
295 300 305	
gtc cag aga gtt tgc ctc cca gac tca tct ata aag ttg cca cct aaa Val Gln Arg Val Cys Leu Pro Asp Ser Ser Ile Lys Leu Pro Pro Lys	1016
310 315 320	
aca agt gtg ttc gtc aca gga ttt gga tcc att gta gat gat gga cct Thr Ser Val Phe Val Thr Gly Phe Gly Ser Ile Val Asp Asp Gly Pro	1064
325 330 335 340	
ata caa aat aca ctt cgg caa gcc aga gtg gaa acc ata agc act gat Ile Gln Asn Thr Leu Arg Gln Ala Arg Val Glu Thr Ile Ser Thr Asp	1112
345 350 355	
gtg tgt aac aga aag gat gtg tat gat ggc ctg ata act cca gga atg Val Cys Asn Arg Lys Asp Val Tyr Asp Gly Leu Ile Thr Pro Gly Met	1160
360 365 370	
tta tgt gct gga ttc atg gaa gga aaa ata gat gca tgt aag gga gat Leu Cys Ala Gly Phe Met Glu Gly Lys Ile Asp Ala Cys Lys Gly Asp	1208
375 380 385	
tct ggt gga cct ctg gtt tat gat aat cat gac atc tgg tac att gta Ser Gly Gly Pro Leu Val Tyr Asp Asn His Asp Ile Trp Tyr Ile Val	1256
390 395 400	
ggt ata gta agt tgg gga caa tca tgt gca ctt ccc aaa aaa cct gga Gly Ile Val Ser Trp Gly Gln Ser Cys Ala Leu Pro Lys Lys Pro Gly	1304
405 410 415 420	
gtc tac acc aga gta act aag tat cga gat tgg att gcc tca aag act Val Tyr Thr Arg Val Thr Lys Tyr Arg Asp Trp Ile Ala Ser Lys Thr	1352
425 430 435	
ggt atg tag tgtggattgt ccatgaggtta tacacatggc acacagagct Gly Met *	1401
gatactcctg cgtatTTTTgt attgTTTTaaa ttcattttact ttggatttagt gctttttgcta	1461
gatgtcaaga agcccttcag acccagacaa atctaataac ctgaggtggc ctttacatac	1521
gtaggaccaa accctctcta ccatgaggga agaagacaca gcaaatgaca gacagcacct	1581
attccttact cacaagggaa actgcttggtg ataacttcta ataagataaa taagtggttt	1641
ccctcaattg aagacaggaa catcattttc cacaggatat gaagagctgc cagtaatgcc	1701
aaaatcttac ctcatataat acctggagca tgtgagattc ttctagttaa aaagaacagt	1761
cttccttgaa gactcagggc ttcaacattc tagaactgat aagtggacct tcagtgtgca	1821
agaatggaga agcatgggat ttgcattatg acttgaactg ggcttatatc taataatata	1881
gagcactatc actaacctca acagttgaca ttttaaaagt ttttaaatgt atctgaactt	1941
gctgttaaca cagtgttata actcaagcac tagcttcagg aagcatgttg tgttgtttaag	2001
aagcttttct gatttattct ttaacagcat cttgccatct atatgttagt agcagttggc	2061
ccagaaagga caaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	2100

<400>	14															
Met	Met	Tyr	Thr	Pro	Val	Glu	Phe	Ser	Glu	Ala	Glu	Phe	Ser	Arg	Ala	
1				5					10					15		
Glu	Tyr	Gln	Arg	Lys	Gln	Gln	Phe	Trp	Asp	Ser	Val	Arg	Leu	Ala	Leu	
			20					25					30			
Phe	Thr	Leu	Ala	Ile	Val	Ala	Ile	Ile	Gly	Ile	Ala	Ile	Gly	Ile	Val	
		35					40					45				
Thr	His	Phe	Val	Val	Glu	Asp	Asp	Lys	Ser	Phe	Tyr	Tyr	Leu	Ala	Ser	
	50					55					60					
Phe	Lys	Val	Thr	Asn	Ile	Lys	Tyr	Lys	Glu	Asn	Tyr	Gly	Ile	Arg	Ser	
65					70					75					80	
Ser	Arg	Glu	Phe	Ile	Glu	Arg	Ser	His	Gln	Ile	Glu	Arg	Met	Met	Ser	
				85					90					95		
Arg	Ile	Phe	Arg	His	Ser	Ser	Val	Gly	Gly	Arg	Phe	Ile	Lys	Ser	His	
			100					105					110			
Val	Ile	Lys	Leu	Ser	Pro	Asp	Glu	Gln	Gly	Val	Asp	Ile	Leu	Ile	Val	
		115					120					125				
Leu	Ile	Phe	Arg	Tyr	Pro	Ser	Thr	Asp	Ser	Ala	Glu	Gln	Ile	Lys	Lys	
	130					135					140					
Lys	Ile	Glu	Lys	Ala	Leu	Tyr	Gln	Ser	Leu	Lys	Thr	Lys	Gln	Leu	Ser	
145					150					155					160	
Leu	Thr	Ile	Asn	Lys	Pro	Ser	Phe	Arg	Leu	Thr	Pro	Ile	Asp	Ser	Lys	
			165						170				175			
Lys	Met	Arg	Asn	Leu	Leu	Asn	Ser	Arg	Cys	Gly	Ile	Arg	Met	Thr	Ser	
			180					185					190			
Ser	Asn	Met	Pro	Leu	Pro	Ala	Ser	Ser	Ser	Thr	Gln	Arg	Ile	Val	Gln	
		195					200					205				
Gly	Arg	Glu	Thr	Ala	Met	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Ala	Ser	Leu	
	210					215					220					
Gln	Leu	Ile	Gly	Ser	Gly	His	Gln	Cys	Gly	Ala	Ser	Leu	Ile	Ser	Asn	
225					230					235					240	
Thr	Trp	Leu	Leu	Thr	Ala	Ala	His	Cys	Phe	Trp	Lys	Asn	Lys	Asp	Pro	
			245						250				255			
Thr	Gln	Trp	Ile	Ala	Thr	Phe	Gly	Ala	Thr	Ile	Thr	Pro	Pro	Ala	Val	
			260				265						270			
Lys	Arg	Asn	Val	Arg	Lys	Ile	Ile	Leu	His	Glu	Asn	Tyr	His	Arg	Glu	
		275					280				285					
Thr	Asn	Glu	Asn	Asp	Ile	Ala	Leu	Val	Gln	Leu	Ser	Thr	Gly	Val	Glu	
	290					295					300					
Phe	Ser	Asn	Ile	Val	Gln	Arg	Val	Cys	Leu	Pro	Asp	Ser	Ser	Ile	Lys	
305					310					315					320	
Leu	Pro</															

-29-

435

<210> 15  
 <211> 702  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> CDS  
 <222> (1)...(702)  
 <223> Nucleotide sequence encoding MTSP-7 Protease Domain

<400> 15  
 att gtc caa gga agg gaa aca gct atg gaa ggg gaa tgg cca tgg cag 48  
 Ile Val Gln Gly Arg Glu Thr Ala Met Glu Gly Glu Trp Pro Trp Gln  
 1 5 10 15

gcc agc ctc cag ctc ata ggg tca ggc cat cag tgt gga gcc agc ctc 96  
 Ala Ser Leu Gln Leu Ile Gly Ser Gly His Gln Cys Gly Ala Ser Leu  
 20 25 30

atc agt aac aca tgg ctg ctc aca gca gct cac tgc ttt tgg aaa aat 144  
 Ile Ser Asn Thr Trp Leu Leu Thr Ala Ala His Cys Phe Trp Lys Asn  
 35 40 45

aaa gac cca act caa tgg att gct act ttt ggt gca act ata aca cca 192  
 Lys Asp Pro Thr Gln Trp Ile Ala Thr Phe Gly Ala Thr Ile Thr Pro  
 50 55 60

ccc gca gtg aaa cga aat gtg agg aaa att att ctt cat gag aat tac 240  
 Pro Ala Val Lys Arg Asn Val Arg Lys Ile Ile Leu His Glu Asn Tyr  
 65 70 75 80

cat aga gaa aca aat gaa aat gac att gct ttg gtt cag ctc tct act 288  
 His Arg Glu Thr Asn Glu Asn Asp Ile Ala Leu Val Gln Leu Ser Thr  
 85 90 95

gga gtt gag ttt tca aat ata gtc cag aga gtt tgc ctc cca gac tca 336  
 Gly Val Glu Phe Ser Asn Ile Val Gln Arg Val Cys Leu Pro Asp Ser  
 100 105 110

tct ata aag ttg cca cct aaa aca agt gtg ttc gtc aca gga ttt gga 384  
 Ser Ile Lys Leu Pro Pro Lys Thr Ser Val Phe Val Thr Gly Phe Gly  
 115 120 125

tcc att gta gat gat gga cct ata caa aat aca ctt cgg caa gcc aga 432  
 Ser Ile Val Asp Asp Gly Pro Ile Gln Asn Thr Leu Arg Gln Ala Arg  
 130 135 140

gtg gaa acc ata agc act gat gtg tgt aac aga aag gat gtg tat gat 480  
 Val Glu Thr Ile Ser Thr Asp Val Cys Asn Arg Lys Asp Val Tyr Asp  
 145 150 155 160

ggc ctg ata act cca gga atg tta tgt gct gga ttc atg gaa gga aaa 528  
 Gly Leu Ile Thr Pro Gly Met Leu Cys Ala Gly Phe Met Glu Gly Lys  
 165 170 175

ata gat gca tgt aag gga gat tct ggt gga cct ctg gtt tat gat aat 576  
 Ile Asp Ala Cys Lys Gly Asp Ser Gly Gly Pro Leu Val Tyr Asp Asn  
 180 185 190



-30-

```

cat gac atc tgg tac att gta ggt ata gta agt tgg gga caa tca tgt      624
His Asp Ile Trp Tyr Ile Val Gly Ile Val Ser Trp Gly Gln Ser Cys
      195                200                205

```

```

gca ctt ccc aaa aaa cct gga gtc tac acc aga gta act aag tat cga      672
Ala Leu Pro Lys Lys Pro Gly Val Tyr Thr Arg Val Thr Lys Tyr Arg
      210                215                220

```

```

gat tgg att gcc tca aag act ggt atg tag      702
Asp Trp Ile Ala Ser Lys Thr Gly Met *
      225                230

```

<210> 16  
 <211> 233  
 <212> PRT  
 <213> Homo sapien

```

<400> 16
Ile Val Gln Gly Arg Glu Thr Ala Met Glu Gly Glu Trp Pro Trp Gln
  1      5      10      15
Ala Ser Leu Gln Leu Ile Gly Ser Gly His Gln Cys Gly Ala Ser Leu
      20      25      30
Ile Ser Asn Thr Trp Leu Leu Thr Ala Ala His Cys Phe Trp Lys Asn
      35      40      45
Lys Asp Pro Thr Gln Trp Ile Ala Thr Phe Gly Ala Thr Ile Thr Pro
      50      55      60
Pro Ala Val Lys Arg Asn Val Arg Lys Ile Ile Leu His Glu Asn Tyr
      65      70      75      80
His Arg Glu Thr Asn Glu Asn Asp Ile Ala Leu Val Gln Leu Ser Thr
      85      90      95
Gly Val Glu Phe Ser Asn Ile Val Gln Arg Val Cys Leu Pro Asp Ser
      100      105      110
Ser Ile Lys Leu Pro Pro Lys Thr Ser Val Phe Val Thr Gly Phe Gly
      115      120      125
Ser Ile Val Asp Asp Gly Pro Ile Gln Asn Thr Leu Arg Gln Ala Arg
      130      135      140
Val Glu Thr Ile Ser Thr Asp Val Cys Asn Arg Lys Asp Val Tyr Asp
      145      150      155      160
Gly Leu Ile Thr Pro Gly Met Leu Cys Ala Gly Phe Met Glu Gly Lys
      165      170      175
Ile Asp Ala Cys Lys Gly Asp Ser Gly Gly Pro Leu Val Tyr Asp Asn
      180      185      190
His Asp Ile Trp Tyr Ile Val Gly Ile Val Ser Trp Gly Gln Ser Cys
      195      200      205
Ala Leu Pro Lys Lys Pro Gly Val Tyr Thr Arg Val Thr Lys Tyr Arg
      210      215      220
Asp Trp Ile Ala Ser Lys Thr Gly Met
      225      230

```

<210> 17  
 <211> 777  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <221> CDS  
 <222> (1)...(729)  
 <223> Nucleotide sequence encoding MTSP9, including  
 protease domain (31-729)

-31-

```

<400> 17
aaa cga gtt gtt cca tta aac gtc aac aga ata gca tct gga gtc att 48
Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile
1 5 10 15

gca ccc aag gcg gcc tgg cct tgg caa gct tcc ctt cag tat gat aac 96
Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn
20 25 30

atc cat cag tgt ggg gcc acc ttg att agt aac aca tgg ctt gtc act 144
Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr
35 40 45

gca gca cac tgc ttc cag aag tat aaa aat cca cat caa tgg act gtt 192
Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val
50 55 60

agt ttt gga aca aaa atc aac cct ccc tta atg aaa aga aat gtc aga 240
Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg
65 70 75 80

aga ttt att atc cat gag aag tac cgc tct gca gca aga gag tac gac 288
Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp
85 90 95

att gct gtt gtg cag gtc tct tcc aga gtc acc ttt tcg gat gac ata 336
Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile
100 105 110

cgc cgg att tgt ttg cca gaa gcc tct gca tcc ttc caa cca aat ttg 384
Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu
115 120 125

act gtc cac atc aca gga ttt gga gca ctt tac tat ggt ggg gaa tcc 432
Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser
130 135 140

caa aat gat ctc cga gaa gcc aga gtg aaa atc ata agt gac gat gtc 480
Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val
145 150 155 160

tgc aag caa cca cag gtg tat ggc aat gat ata aaa cct gga atg ttc 528
Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe
165 170 175

tgt gcc gga tat atg gaa gga att tat gat gcc tgc agg ggt gat tct 576
Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser
180 185 190

ggg gga cct tta gtc aca agg gat ctg aaa gat acg tgg tat ctc att 624
Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile
195 200 205

gga att gta agc tgg gga gat aac tgt ggt caa aag gac aag cct gga 672
Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly
210 215 220

gtc tac aca caa gtg act tat tac cga aac tgg att gct tca aaa aca 720
Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr
225 230 235 240

```

-32-

ggc atc taa ttcacgataa aagttaaaca aagaaagctg tatgcaggtc atatatgc 777  
Gly Ile

<210> 18  
<211> 242  
<212> PRT  
<213> Homo Sapien

<220>  
<221> SITE  
<222> (11)...(242)  
<223> MTSP9 protease domain

<400> 18  
Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile  
1 5 10 15  
Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn  
20 25 30  
Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr  
35 40 45  
Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val  
50 55 60  
Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg  
65 70 75 80  
Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp  
85 90 95  
Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile  
100 105 110  
Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu  
115 120 125  
Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser  
130 135 140  
Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val  
145 150 155 160  
Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe  
165 170 175  
Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser  
180 185 190  
Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile  
195 200 205  
Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly  
210 215 220  
Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr  
225 230 235 240  
Gly Ile

-33-

<210> 19 MTSP12  
 <211> 3316  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <221> CDS  
 <222> (1)...(3282)  
 <223> Nucleotide sequence encoding MTSP12, including  
 MTSP12-PD1, MTSP12-PD2, and MTSP12-PD3 protease  
 domains

<400> 19  
 atg gag ccc act gtg gct aac gta cac ctc gtg ccc agg aca acc aag 48  
 Met Glu Pro Thr Val Ala Asn Val His Leu Val Pro Arg Thr Thr Lys  
 1 5 10 15  
 gaa gtc ccc gct ctg gat gcc gcg tgc tgt cga gcg gcc acc att ggc 96  
 Glu Val Pro Ala Leu Asp Ala Ala Cys Cys Arg Ala Ala Thr Ile Gly  
 20 25 30  
 gtg gtg gcc acc agc ctt gtc gtc ctc acc ctg gga gtc ctt ttg gcc 144  
 Val Val Ala Thr Ser Leu Val Val Leu Thr Leu Gly Val Leu Leu Ala  
 35 40 45  
 ttc ctc tct aca cag ggc ttc cac gtg gac cac acg gcc gag ctg cgg 192  
 Phe Leu Ser Thr Gln Gly Phe His Val Asp His Thr Ala Glu Leu Arg  
 50 55 60  
 gga atc cgg tgg acc agc agt ttg cgg cgg gag acc tcg gac tat cac 240  
 Gly Ile Arg Trp Thr Ser Ser Leu Arg Arg Glu Thr Ser Asp Tyr His  
 65 70 75 80  
 cgc acg ctg acg ccc acc ctg gag gca ctg ttt gta agt agt ttt cag 288  
 Arg Thr Leu Thr Pro Thr Leu Glu Ala Leu Phe Val Ser Ser Phe Gln  
 85 90 95  
 aag aca gag tta gag gca agc tgc gtg ggt tgc tcg gta ctg aat tat 336  
 Lys Thr Glu Leu Glu Ala Ser Cys Val Gly Cys Ser Val Leu Asn Tyr  
 100 105 110  
 agg gat ggg aac tcc agt gtc ctc gta cat ttc cag ctg cac ttt ctg 384  
 Arg Asp Gly Asn Ser Ser Val Leu Val His Phe Gln Leu His Phe Leu  
 115 120 125  
 ctg cga ccc ctc cag acg ctg agc ctg ggc ctg gag gag gag cta ttg 432  
 Leu Arg Pro Leu Gln Thr Leu Ser Leu Gly Leu Glu Glu Glu Leu Leu  
 130 135 140  
 cag cga ggg atc cgg gca agg ctg cgg gag cac ggc atc tcc ctg gct 480  
 Gln Arg Gly Ile Arg Ala Arg Leu Arg Glu His Gly Ile Ser Leu Ala  
 145 150 155 160  
 gcc tat ggc aca att gtg tcg gct gag ctc aca ggg aga cat aag ggg 528  
 Ala Tyr Gly Thr Ile Val Ser Ala Glu Leu Thr Gly Arg His Lys Gly  
 165 170 175  
 ccc ttg gca gaa aga gac ttc aaa tca ggc cgc tgt cca ggg aac tcc 576  
 Pro Leu Ala Glu Arg Asp Phe Lys Ser Gly Arg Cys Pro Gly Asn Ser  
 180 185 190

-34-

ttt tcc tgc ggg aac agc cag tgt gtg acc aag gtg aac ccg gag tgt	624
Phe Ser Cys Gly Asn Ser Gln Cys Val Thr Lys Val Asn Pro Glu Cys	
195 200 205	
gac gac cag gag gac tgc tcc gat ggg tcc gac gag gcg cac tgc gag	672
Asp Asp Gln Glu Asp Cys Ser Asp Gly Ser Asp Glu Ala His Cys Glu	
210 215 220	
tgt ggc ttg cag cct gcc tgg agg atg gcc ggc agg atc gtg ggc ggc	720
Cys Gly Leu Gln Pro Ala Trp Arg Met Ala Gly Arg Ile Val Gly Gly	
225 230 235 240	
atg gaa gca tcc ccg ggg gag ttt ccg tgg caa gcc agc ctt cga gag	768
Met Glu Ala Ser Pro Gly Glu Phe Pro Trp Gln Ala Ser Leu Arg Glu	
245 250 255	
aac aag gag cac ttc tgt ggg gcc gcc atc atc aac gcc agg tgg ctg	816
Asn Lys Glu His Phe Cys Gly Ala Ala Ile Ile Asn Ala Arg Trp Leu	
260 265 270	
gtg tct gct gct cac tgc ttc aat gag ttc caa gac ccg acg aag tgg	864
Val Ser Ala Ala His Cys Phe Asn Glu Phe Gln Asp Pro Thr Lys Trp	
275 280 285	
gtg gcc tac gtg ggt gcg acc tac ctc agc ggc tgc gag gcc agc acc	912
Val Ala Tyr Val Gly Ala Thr Tyr Leu Ser Gly Ser Glu Ala Ser Thr	
290 295 300	
gtg cgg gcc cag gtg gtc cag atc gtc aag cac ccc ctg tac aac gcg	960
Val Arg Ala Gln Val Val Gln Ile Val Lys His Pro Leu Tyr Asn Ala	
305 310 315 320	
gac acg gcc gac ttt gac gtg gct gtg ctg gag ctg acc agc cct ctg	1008
Asp Thr Ala Asp Phe Asp Val Ala Val Leu Glu Leu Thr Ser Pro Leu	
325 330 335	
cct ttc ggc cgg cac atc cag ccc gtg tgc ctc ccg gct gcc aca cac	1056
Pro Phe Gly Arg His Ile Gln Pro Val Cys Leu Pro Ala Ala Thr His	
340 345 350	
atc ttc cca ccc agc aag aag tgc ctg atc tca ggc tgg ggc tac ctc	1104
Ile Phe Pro Pro Ser Lys Lys Cys Leu Ile Ser Gly Trp Gly Tyr Leu	
355 360 365	
aag gag gac ttc ctg gtc aag cca ggg gtg ctg cag aaa gcc act gtg	1152
Lys Glu Asp Phe Leu Val Lys Pro Gly Val Leu Gln Lys Ala Thr Val	
370 375 380	
gag ctg ctg gac cag gca ctg tgt gcc agc ttg tac ggc cat tca ctc	1200
Glu Leu Leu Asp Gln Ala Leu Cys Ala Ser Leu Tyr Gly His Ser Leu	
385 390 395 400	
act gac agg atg gtg tgc gct ggc tac ctg gac ggg aag gtg gac tcc	1248
Thr Asp Arg Met Val Cys Ala Gly Tyr Leu Asp Gly Lys Val Asp Ser	
405 410 415	
tgc cag ggt gac tca gga gga ccc ctg gtc tgc gag gag ccc tct ggc	1296
Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Glu Glu Pro Ser Gly	
420 425 430	
cgg ttc tct ctg gct ggc atc gtg agc tgg gga atc ggg tgt gcg gaa	1344

-35-

Arg	Phe	Ser	Leu	Ala	Gly	Ile	Val	Ser	Trp	Gly	Ile	Gly	Cys	Ala	Glu		
		435					440					445					
gcc	cgg	cgt	cca	ggg	gtc	tat	gcc	cga	gtc	acc	agg	cta	cgt	gac	tgg		1392
Ala	Arg	Arg	Pro	Gly	Val	Tyr	Ala	Arg	Val	Thr	Arg	Leu	Arg	Asp	Trp		
	450					455					460						
atc	ctg	gag	gcc	acc	acc	aaa	gcc	agc	atg	cct	ctg	gcc	ccc	acc	atg		1440
Ile	Leu	Glu	Ala	Thr	Thr	Lys	Ala	Ser	Met	Pro	Leu	Ala	Pro	Thr	Met		
465					470					475					480		
gct	cct	gcc	cct	gcc	gcc	ccc	agc	aca	gcc	tgg	ccc	acc	agt	cct	gag		1488
Ala	Pro	Ala	Pro	Ala	Ala	Pro	Ser	Thr	Ala	Trp	Pro	Thr	Ser	Pro	Glu		
				485					490					495			
u																	
agc	cct	gtt	gtc	agc	acc	ccc	acc	aaa	tcg	atg	cag	gcc	ctc	agt	acc		1536
Ser	Pro	Val	Val	Ser	Thr	Pro	Thr	Lys	Ser	Met	Gln	Ala	Leu	Ser	Thr		
		500						505					510				
gtg	cct	ctt	gac	tgg	gtc	acc	gtt	cct	aag	cta	caa	gaa	tgt	ggg	gcc		1584
Val	Pro	Leu	Asp	Trp	Val	Thr	Val	Pro	Lys	Leu	Gln	Glu	Cys	Gly	Ala		
		515					520					525					
agg	cct	gca	atg	gag	aag	ccc	acc	cgg	gtc	gtg	ggc	ggg	ttc	gga	gct		1632
Arg	Pro	Ala	Met	Glu	Lys	Pro	Thr	Arg	Val	Val	Gly	Gly	Phe	Gly	Ala		
	530					535					540						
gcc	tcc	ggg	gag	gtg	ccc	tgg	cag	gtc	agc	ctg	aag	gaa	ggg	tcc	cgg		1680
Ala	Ser	Gly	Glu	Val	Pro	Trp	Gln	Val	Ser	Leu	Lys	Glu	Gly	Ser	Arg		
545					550					555					560		
cac	ttc	tgc	gga	gca	act	gtg	gtg	ggg	gac	cgc	tgg	ctg	ctg	tct	gcc		1728
His	Phe	Cys	Gly	Ala	Thr	Val	Val	Gly	Asp	Arg	Trp	Leu	Leu	Ser	Ala		
				565				570						575			
gcc	cac	tgc	ttc	aac	cac	acg	aag	gtg	gag	cag	gtt	cgg	gcc	cac	ctg		1776
Ala	His	Cys	Phe	Asn	His	Thr	Lys	Val	Glu	Gln	Val	Arg	Ala	His	Leu		
			580					585					590				
ggc	act	gcg	tcc	ctc	ctg	ggc	ctg	ggc	ggg	agc	ccg	gtg	aag	atc	ggg		1824
Gly	Thr	Ala	Ser	Leu	Leu	Gly	Leu	Gly	Gly	Ser	Pro	Val	Lys	Ile	Gly		
		595				600						605					
ctg	cgg	cgg	gta	gtg	ctg	cac	ccc	ctc	tac	aac	cct	ggc	atc	ctg	gac		1872
Leu	Arg	Arg	Val	Val	Leu	His	Pro	Leu	Tyr	Asn	Pro	Gly	Ile	Leu	Asp		
	610					615					620						
ttc	gac	ctg	gct	gtc	ctg	gag	ctg	gcc	agc	ccc	ctg	gcc	ttc	aac	aaa		1920
Phe	Asp	Leu	Ala	Val	Leu	Glu	Leu	Ala	Ser	Pro	Leu	Ala	Phe	Asn	Lys		
625					630					635					640		
tac	atc	cag	cct	gtc	tgc	ctg	ccc	ctg	gcc	atc	cgg	aag	ttc	cct	gtg		1968
Tyr	Ile	Gln	Pro	Val	Cys	Leu	Pro	Leu	Ala	Ile	Arg	Lys	Phe	Pro	Val		
				645				650						655			
ggc	cgg	aag	tgc	atg	atc	tcc	gga	tgg	gga	aat	acg	cag	gaa	gga	aat		2016
Gly	Arg	Lys	Cys	Met	Ile	Ser	Gly	Trp	Gly	Asn	Thr	Gln	Glu	Gly	Asn		
			660				665						670				
gcc	acc	aag	ccc	gag	ctc	ctg	cag	aag	gcg	tcc	gtg	ggc	atc	ata	gac		2064
Ala	Thr	Lys	Pro	Glu	Leu	Leu	Gln	Lys	Ala	Ser	Val	Gly	Ile	Ile	Asp		

-36-

675	680	685	
cag aaa acc tgt agt gtg ctc tac aac ttc tcc ctc aca gac cgc atg Gln Lys Thr Cys Ser Val Leu Tyr Asn Phe Ser Leu Thr Asp Arg Met 690 695 700			2112
atc tgc gca ggc ttc ctg gaa ggc aaa gtc gac tcc tgc cag ggt gac Ile Cys Ala Gly Phe Leu Glu Gly Lys Val Asp Ser Cys Gln Gly Asp 705 710 715 720			2160
tct ggg ggc ccc ctg gcc tgc gag gag gcc cct ggc gtg ttt tat ctg Ser Gly Gly Pro Leu Ala Cys Glu Glu Ala Pro Gly Val Phe Tyr Leu 725 730 735			2208
gca ggg atc gtg agc tgg ggt att ggc tgc gct cag gtt aag aag ccg Ala Gly Ile Val Ser Trp Gly Ile Gly Cys Ala Gln Val Lys Lys Pro 740 745 750			2256
ggc gtg tac acg cgc atc acc agg cta aag ggc tgg atc ctg gag atc Gly Val Tyr Thr Arg Ile Thr Arg Leu Lys Gly Trp Ile Leu Glu Ile 755 760 765			2304
atg tcc tcc cag ccc ctt ccc atg tct ccc ccc tcc acc aca agg atg Met Ser Ser Gln Pro Leu Pro Met Ser Pro Pro Ser Thr Thr Arg Met 770 775 780			2352
ctg gcc acc acc agc ccc agg acg aca gct ggc ctc aca gtc ccg ggg Leu Ala Thr Thr Ser Pro Arg Thr Thr Ala Gly Leu Thr Val Pro Gly 785 790 795 800			2400
gcc aca ccc agc aga ccc acc cct ggg gct gcc agc agg gtg acg ggc Ala Thr Pro Ser Arg Pro Thr Pro Gly Ala Ala Ser Arg Val Thr Gly 805 810 815			2448
caa cct gcc aac tca acc tta tct gcc gtg agc acc act gct agg gga Gln Pro Ala Asn Ser Thr Leu Ser Ala Val Ser Thr Thr Ala Arg Gly 820 825 830			2496
cag acg cca ttt cca gac gcc ccg gag gcc acc aca cac acc cag cta Gln Thr Pro Phe Pro Asp Ala Pro Glu Ala Thr Thr His Thr Gln Leu 835 840 845			2544
cca gac tgt ggc ctg gcg ccg gcc gcg ctc acc agg att gtg ggc ggc Pro Asp Cys Gly Leu Ala Pro Ala Ala Leu Thr Arg Ile Val Gly Gly 850 855 860			2592
agc gca gcg ggc cgt ggg gag tgg ccg tgg cag gtg ggc ctg tgg ctg Ser Ala Ala Gly Arg Gly Glu Trp Pro Trp Gln Val Gly Leu Trp Leu 865 870 875 880			2640
cgg cgc cgg gaa cac cgt tgc ggg gcc gtg ctg gtg gca gag agg tgg Arg Arg Arg Glu His Arg Cys Gly Ala Val Leu Val Ala Glu Arg Trp 885 890 895			2688
ctg ctg tcc gcg gcg cac tgc ttc gac gtc tac ggg gac ccc aag cag Leu Leu Ser Ala Ala His Cys Phe Asp Val Tyr Gly Asp Pro Lys Gln 900 905 910			2736
tgg gcg gcc ttc cta ggc acg ccg ttc ctg agc ggc gcg gag ggg cag Trp Ala Ala Phe Leu Gly Thr Pro Phe Leu Ser Gly Ala Glu Gly Gln 915 920 925			2784

-37-

ctg gag cgc gtg gcg cgc atc tac aag cac ccg ttc tac aat ctc tac	2832
Leu Glu Arg Val Ala Arg Ile Tyr Lys His Pro Phe Tyr Asn Leu Tyr	
930 935 940	
acg ctc gac tac gac gtg gcg ctt ctg gag ctg gcg ggg ccg gtg cgt	2880
Thr Leu Asp Tyr Asp Val Ala Leu Leu Glu Leu Ala Gly Pro Val Arg	
945 950 955 960	
cgc agc cgc ctg gtg cgt ccc atc tgc ctg ccc gag ccc gcg ccg cga	2928
Arg Ser Arg Leu Val Arg Pro Ile Cys Leu Pro Glu Pro Ala Pro Arg	
965 970 975	
ccc ccg gac ggc acg cgc tgc gtc atc acc ggc tgg ggc tcg gtg cgc	2976
Pro Pro Asp Gly Thr Arg Cys Val Ile Thr Gly Trp Gly Ser Val Arg	
980 985 990	
gaa gga ggc tcc atg gcg cgg cag ctg cag aag gcg gcc gtg cgc ctc	3024
Glu Gly Gly Ser Met Ala Arg Gln Leu Gln Lys Ala Ala Val Arg Leu	
995 1000 1005	
ctc agc gag cag acc tgc cgc cgc ttc tac cca gtg cag atc agc agc	3072
Leu Ser Glu Gln Thr Cys Arg Arg Phe Tyr Pro Val Gln Ile Ser Ser	
1010 1015 1020	
cgc atg ctg tgt gcc ggc ttc ccg cag ggt ggc gtg gac agc tgc tcg	3120
Arg Met Leu Cys Ala Gly Phe Pro Gln Gly Gly Val Asp Ser Cys Ser	
1025 1030 1035 1040	
ggt gac gct ggg gga ccc ctg gcc tgc agg gag ccc tct gga cgg tgg	3168
Gly Asp Ala Gly Gly Pro Leu Ala Cys Arg Glu Pro Ser Gly Arg Trp	
1045 1050 1055	
gtg cta act ggg gtc act agc tgg ggc tat ggc tgt ggc cgg ccc cac	3216
Val Leu Thr Gly Val Thr Ser Trp Gly Tyr Gly Cys Gly Arg Pro His	
1060 1065 1070	
ttc cca ggt gtc tat acc cgg gtg gca gct gtg aga ggc tgg ata gga	3264
Phe Pro Gly Val Tyr Thr Arg Val Ala Ala Val Arg Gly Trp Ile Gly	
1075 1080 1085	
cag cac atc cag gag tga ccaccacgtg actgcccagg ccgagactct	3312
Gln His Ile Gln Glu *	
1090	
acgt	3316
<210> 20	
<211> 1093	
<212> PRT	
<213> Homo Sapien	
<400> 20	
Met Glu Pro Thr Val Ala Asn Val His Leu Val Pro Arg Thr Thr Lys	
1 5 10 15	
Glu Val Pro Ala Leu Asp Ala Ala Cys Cys Arg Ala Ala Thr Ile Gly	
20 25 30	
Val Val Ala Thr Ser Leu Val Val Leu Thr Leu Gly Val Leu Leu Ala	
35 40 45	
Phe Leu Ser Thr Gln Gly Phe His Val Asp His Thr Ala Glu Leu Arg	
50 55 60	



-38-

Gly	Ile	Arg	Trp	Thr	Ser	Ser	Leu	Arg	Arg	Glu	Thr	Ser	Asp	Tyr	His
65					70					75					80
Arg	Thr	Leu	Thr	Pro	Thr	Leu	Glu	Ala	Leu	Phe	Val	Ser	Ser	Phe	Gln
				85					90					95	
Lys	Thr	Glu	Leu	Glu	Ala	Ser	Cys	Val	Gly	Cys	Ser	Val	Leu	Asn	Tyr
			100				105						110		
Arg	Asp	Gly	Asn	Ser	Ser	Val	Leu	Val	His	Phe	Gln	Leu	His	Phe	Leu
		115					120					125			
Leu	Arg	Pro	Leu	Gln	Thr	Leu	Ser	Leu	Gly	Leu	Glu	Glu	Glu	Leu	Leu
		130				135					140				
Gln	Arg	Gly	Ile	Arg	Ala	Arg	Leu	Arg	Glu	His	Gly	Ile	Ser	Leu	Ala
145					150					155					160
Ala	Tyr	Gly	Thr	Ile	Val	Ser	Ala	Glu	Leu	Thr	Gly	Arg	His	Lys	Gly
				165					170					175	
Pro	Leu	Ala	Glu	Arg	Asp	Phe	Lys	Ser	Gly	Arg	Cys	Pro	Gly	Asn	Ser
			180				185						190		
Phe	Ser	Cys	Gly	Asn	Ser	Gln	Cys	Val	Thr	Lys	Val	Asn	Pro	Glu	Cys
		195					200					205			
Asp	Asp	Gln	Glu	Asp	Cys	Ser	Asp	Gly	Ser	Asp	Glu	Ala	His	Cys	Glu
	210					215					220				
Cys	Gly	Leu	Gln	Pro	Ala	Trp	Arg	Met	Ala	Gly	Arg	Ile	Val	Gly	Gly
225					230					235					240
Met	Glu	Ala	Ser	Pro	Gly	Glu	Phe	Pro	Trp	Gln	Ala	Ser	Leu	Arg	Glu
				245					250					255	
Asn	Lys	Glu	His	Phe	Cys	Gly	Ala	Ala	Ile	Ile	Asn	Ala	Arg	Trp	Leu
			260				265						270		
Val	Ser	Ala	Ala	His	Cys	Phe	Asn	Glu	Phe	Gln	Asp	Pro	Thr	Lys	Trp
		275					280					285			
Val	Ala	Tyr	Val	Gly	Ala	Thr	Tyr	Leu	Ser	Gly	Ser	Glu	Ala	Ser	Thr
	290					295					300				
Val	Arg	Ala	Gln	Val	Val	Gln	Ile	Val	Lys	His	Pro	Leu	Tyr	Asn	Ala
305					310					315					320
Asp	Thr	Ala	Asp	Phe	Asp	Val	Ala	Val	Leu	Glu	Leu	Thr	Ser	Pro	Leu
				325					330					335	
Pro	Phe	Gly	Arg	His	Ile	Gln	Pro	Val	Cys	Leu	Pro	Ala	Ala	Thr	His
			340					345					350		
Ile	Phe	Pro	Pro	Ser	Lys	Lys	Cys	Leu	Ile	Ser	Gly	Trp	Gly	Tyr	Leu
		355					360					365			
Lys	Glu	Asp	Phe	Leu	Val	Lys	Pro	Gly	Val	Leu	Gln	Lys	Ala	Thr	Val
	370					375					380				
Glu	Leu	Leu	Asp	Gln	Ala	Leu	Cys	Ala	Ser	Leu	Tyr	Gly	His	Ser	Leu
385					390					395					400
Thr	Asp	Arg	Met	Val	Cys	Ala	Gly	Tyr	Leu	Asp	Gly	Lys	Val	Asp	Ser
			405						410					415	
Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys	Glu	Glu	Pro	Ser	Gly
			420					425					430		
Arg	Phe	Ser	Leu	Ala	Gly	Ile	Val	Ser	Trp	Gly	Ile	Gly	Cys	Ala	Glu
		435					440					445			
Ala	Arg	Arg	Pro	Gly	Val	Tyr	Ala	Arg	Val	Thr	Arg	Leu	Arg	Asp	Trp
	450					455					460				
Ile	Leu	Glu	Ala	Thr	Thr	Lys	Ala	Ser	Met	Pro	Leu	Ala	Pro	Thr	Met
465					470					475					480
Ala	Pro	Ala	Pro	Ala	Ala	Pro	Ser	Thr	Ala	Trp	Pro	Thr	Ser	Pro	Glu
				485					490					495	
Ser	Pro	Val	Val	Ser	Thr	Pro	Thr	Lys	Ser	Met	Gln	Ala	Leu	Ser	Thr
			500					505					510		
Val	Pro	Leu	Asp	Trp	Val	Thr	Val	Pro	Lys	Leu	Gln	Glu	Cys	Gly	Ala
		515					520					525			
Arg	Pro	Ala	Met	Glu	Lys	Pro	Thr	Arg	Val	Val	Gly	Gly	Phe	Gly	Ala
	530					535					540				
Ala	Ser	Gly	Glu	Val	Pro	Trp	Gln	Val	Ser	Leu	Lys	Glu	Gly	Ser	Arg

-39-

545					550					555				560
His	Phe	Cys	Gly	Ala	Thr	Val	Val	Gly	Asp	Arg	Trp	Leu	Leu	Ser
				565					570					575
Ala	His	Cys	Phe	Asn	His	Thr	Lys	Val	Glu	Gln	Val	Arg	Ala	His
			580					585					590	
Gly	Thr	Ala	Ser	Leu	Leu	Gly	Leu	Gly	Gly	Ser	Pro	Val	Lys	Ile
		595					600					605		
Leu	Arg	Arg	Val	Val	Leu	His	Pro	Leu	Tyr	Asn	Pro	Gly	Ile	Leu
	610				615						620			
Phe	Asp	Leu	Ala	Val	Leu	Glu	Leu	Ala	Ser	Pro	Leu	Ala	Phe	Asn
625					630				635					640
Tyr	Ile	Gln	Pro	Val	Cys	Leu	Pro	Leu	Ala	Ile	Arg	Lys	Phe	Pro
			645						650					655
Gly	Arg	Lys	Cys	Met	Ile	Ser	Gly	Trp	Gly	Asn	Thr	Gln	Glu	Gly
			660					665					670	
Ala	Thr	Lys	Pro	Glu	Leu	Leu	Gln	Lys	Ala	Ser	Val	Gly	Ile	Ile
		675					680					685		
Gln	Lys	Thr	Cys	Ser	Val	Leu	Tyr	Asn	Phe	Ser	Leu	Thr	Asp	Arg
	690					695					700			
Ile	Cys	Ala	Gly	Phe	Leu	Glu	Gly	Lys	Val	Asp	Ser	Cys	Gln	Gly
705					710					715				720
Ser	Gly	Gly	Pro	Leu	Ala	Cys	Glu	Glu	Ala	Pro	Gly	Val	Phe	Tyr
			725						730					735
Ala	Gly	Ile	Val	Ser	Trp	Gly	Ile	Gly	Cys	Ala	Gln	Val	Lys	Lys
			740					745					750	
Gly	Val	Tyr	Thr	Arg	Ile	Thr	Arg	Leu	Lys	Gly	Trp	Ile	Leu	Glu
	755						760					765		
Met	Ser	Ser	Gln	Pro	Leu	Pro	Met	Ser	Pro	Pro	Ser	Thr	Thr	Arg
	770					775					780			
Leu	Ala	Thr	Thr	Ser	Pro	Arg	Thr	Thr	Ala	Gly	Leu	Thr	Val	Pro
785					790				795					800
Ala	Thr	Pro	Ser	Arg	Pro	Thr	Pro	Gly	Ala	Ala	Ser	Arg	Val	Thr
			805						810					815
Gln	Pro	Ala	Asn	Ser	Thr	Leu	Ser	Ala	Val	Ser	Thr	Thr	Ala	Arg
			820					825					830	
Gln	Thr	Pro	Phe	Pro	Asp	Ala	Pro	Glu	Ala	Thr	Thr	His	Thr	Gln
	835						840					845		
Pro	Asp	Cys	Gly	Leu	Ala	Pro	Ala	Ala	Leu	Thr	Arg	Ile	Val	Gly
	850					855					860			
Ser	Ala	Ala	Gly	Arg	Gly	Glu	Trp	Pro	Trp	Gln	Val	Gly	Leu	Trp
865					870				875					880
Arg	Arg	Arg	Glu	His	Arg	Cys	Gly	Ala	Val	Leu	Val	Ala	Glu	Arg
			885						890					895
Leu	Leu	Ser	Ala	Ala	His	Cys	Phe	Asp	Val	Tyr	Gly	Asp	Pro	Lys
			900					905				910		
Trp	Ala	Ala	Phe	Leu	Gly	Thr	Pro	Phe	Leu	Ser	Gly	Ala	Glu	Gly
	915						920					925		
Leu	Glu	Arg	Val	Ala	Arg	Ile	Tyr	Lys	His	Pro	Phe	Tyr	Asn	Leu
	930					935					940			
Thr	Leu	Asp	Tyr	Asp	Val	Ala	Leu	Leu	Glu	Leu	Ala	Gly	Pro	Val
945					950				955					960
Arg	Ser	Arg	Leu	Val	Arg	Pro	Ile	Cys	Leu	Pro	Glu	Pro	Ala	Pro
			965						970					975
Pro	Pro	Asp	Gly	Thr	Arg	Cys	Val	Ile	Thr	Gly	Trp	Gly	Ser	Val
			980					985					990	
Glu	Gly	Gly	Ser	Met	Ala	Arg	Gln	Leu	Gln	Lys	Ala	Ala	Val	Arg
	995						1000					1005		
Leu	Ser	Glu	Gln	Thr	Cys	Arg	Arg	Phe	Tyr	Pro	Val	Gln	Ile	Ser
	1010					1015					1020			
Arg	Met	Leu	Cys	Ala	Gly	Phe	Pro	Gln	Gly	Gly	Val	Asp	Ser	Cys
1025					1030				1035					1040

[illegible]

```
<210> 21
<211> 702
<212> DNA
<213> Homo Sapien
```

```
<220>
<221> CDS
<222> (1)...(699)
<223> Nucleic Acid encoding protease domain of
        endotheliase 1
```

<400> 21																	
agg	atc	gtt	ggt	ggg	aca	gaa	gta	gaa	gag	ggg	gaa	tgg	ccc	tgg	cag		48
Arg	Ile	Val	Gly	Gly	Thr	Glu	Val	Glu	Glu	Gly	Glu	Trp	Pro	Trp	Gln		
1				5					10					15			
gct	agc	ctg	cag	tgg	gat	ggg	agt	cat	cgc	tgt	gga	gca	acc	tta	att		96
Ala	Ser	Leu	Gln	Trp	Asp	Gly	Ser	His	Arg	Cys	Gly	Ala	Thr	Leu	Ile		
			20					25					30				
aat	gcc	aca	tgg	ctt	gtg	agt	gct	gct	cac	tgt	ttt	aca	aca	tat	aag		144
Asn	Ala	Thr	Trp	Leu	Val	Ser	Ala	Ala	His	Cys	Phe	Thr	Thr	Tyr	Lys		
		35					40					45					
aac	cct	gcc	aga	tgg	act	gct	tcc	ttt	gga	gta	aca	ata	aaa	cct	tcg		192
Asn	Pro	Ala	Arg	Trp	Thr	Ala	Ser	Phe	Gly	Val	Thr	Ile	Lys	Pro	Ser		
	50					55					60						
aaa	atg	aaa	cgg	ggt	ctc	cgg	aga	ata	att	gtc	cat	gaa	aaa	tac	aaa		240
Lys	Met	Lys	Arg	Gly	Leu	Arg	Arg	Ile	Ile	Val	His	Glu	Lys	Tyr	Lys		
65				70						75					80		
cac	cca	tca	cat	gac	tat	gat	att	tct	ctt	gca	gag	ctt	tct	agc	cct		288
His	Pro	Ser	His	Asp	Tyr	Asp	Ile	Ser	Leu	Ala	Glu	Leu	Ser	Ser	Pro		
				85				90					95				
gtt	ccc	tac	aca	aat	gca	gta	cat	aga	gtt	tgt	ctc	cct	gat	gca	tcc		336
Val	Pro	Tyr	Thr	Asn	Ala	Val	His	Arg	Val	Cys	Leu	Pro	Asp	Ala	Ser		
			100					105				110					
tat	gag	ttt	caa	cca	ggt	gat	gtg	atg	ttt	gtg	aca	gga	ttt	gga	gca		384
Tyr	Glu	Phe	Gln	Pro	Gly	Asp	Val	Met	Phe	Val	Thr	Gly	Phe	Gly	Ala		
		115					120					125					
ctg	aaa	aat	gat	ggt	tac	agt	caa	aat	cat	ctt	cga	caa	gca	cag	gtg		432
Leu	Lys	Asn	Asp	Gly	Tyr	Ser	Gln	Asn	His	Leu	Arg	Gln	Ala	Gln	Val		
	130					135					140						
act	ctc	ata	gac	gct	aca	act	tgc	aat	gaa	cct	caa	gct	tac	aat	gac		480
Thr	Leu	Ile	Asp	Ala	Thr	Thr	Cys	Asn	Glu	Pro	Gln	Ala	Tyr	Asn	Asp		
145				150						155					160		

-41-

```

gcc ata act cct aga atg tta tgt gct ggc tcc tta gaa gga aaa aca      528
Ala Ile Thr Pro Arg Met Leu Cys Ala Gly Ser Leu Glu Gly Lys Thr
                      165                      170                      175

gat gca tgc cag ggt gac tct gga gga cca ctg gtt agt tca gat gct      576
Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Ser Ser Asp Ala
                      180                      185                      190

aga gat atc tgg tac ctt gct gga ata gtg agc tgg gga gat gaa tgt      624
Arg Asp Ile Trp Tyr Leu Ala Gly Ile Val Ser Trp Gly Asp Glu Cys
                      195                      200                      205

gcg aaa ccc aac aag cct ggt gtt tat act aga gtt acg gcc ttg cgg      672
Ala Lys Pro Asn Lys Pro Gly Val Tyr Thr Arg Val Thr Ala Leu Arg
                      210                      215                      220

gac tgg att act tca aaa act ggt atc taa      702
Asp Trp Ile Thr Ser Lys Thr Gly Ile
225                      230

```

<210> 22  
 <211> 233  
 <212> PRT  
 <213> Homo Sapien

<220>  
 <221> SITE  
 <222> (1)...(233)  
 <223> Protease domain of endotheliase 1

```

<400> 22
Arg Ile Val Gly Gly Thr Glu Val Glu Glu Gly Glu Trp Pro Trp Gln
 1           5           10
Ala Ser Leu Gln Trp Asp Gly Ser His Arg Cys Gly Ala Thr Leu Ile
 20           25           30
Asn Ala Thr Trp Leu Val Ser Ala Ala His Cys Phe Thr Thr Tyr Lys
 35           40           45
Asn Pro Ala Arg Trp Thr Ala Ser Phe Gly Val Thr Ile Lys Pro Ser
 50           55           60
Lys Met Lys Arg Gly Leu Arg Arg Ile Ile Val His Glu Lys Tyr Lys
 65           70           75
His Pro Ser His Asp Tyr Asp Ile Ser Leu Ala Glu Leu Ser Ser Pro
 85           90           95
Val Pro Tyr Thr Asn Ala Val His Arg Val Cys Leu Pro Asp Ala Ser
100          105          110
Tyr Glu Phe Gln Pro Gly Asp Val Met Phe Val Thr Gly Phe Gly Ala
115          120          125
Leu Lys Asn Asp Gly Tyr Ser Gln Asn His Leu Arg Gln Ala Gln Val
130          135          140
Thr Leu Ile Asp Ala Thr Thr Cys Asn Glu Pro Gln Ala Tyr Asn Asp
145          150          155
Ala Ile Thr Pro Arg Met Leu Cys Ala Gly Ser Leu Glu Gly Lys Thr
165          170          175
Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Ser Ser Asp Ala
180          185          190
Arg Asp Ile Trp Tyr Leu Ala Gly Ile Val Ser Trp Gly Asp Glu Cys
195          200          205
Ala Lys Pro Asn Lys Pro Gly Val Tyr Thr Arg Val Thr Ala Leu Arg
210          215          220
Asp Trp Ile Thr Ser Lys Thr Gly Ile

```

-42-

225

230

&lt;210&gt; 23

&lt;211&gt; 1689

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1689)

&lt;223&gt; Nucleic acid encoding Endotheliase 2-S protein

&lt;400&gt; 23

atg gag agg gac agc cac ggg aat gca tct cca gca aga aca cct tca	48
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro Ser	
1 5 10 15	
gct gga gca tct cca gcc cag gca tct cca gct ggg aca cct cca ggc	96
Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro Pro Gly	
20 25 30	
cgg gca tct cca gcc cag gca tct cca gcc cag gca tct cca gct ggg	144
Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly	
35 40 45	
aca cct ccg ggc cgg gca tct cca gcc cag gca tct cca gct ggt aca	192
Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr	
50 55 60	
cct cca ggc cgg gca tct cca gcc cgg gca tct cca gcc cag gca tct	240
Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser	
65 70 75 80	
cca gcc cgg gca tct ccg gct ctg gca tca ctt tcc agg tcc tca tcc	288
Pro Ala Arg Ala Ser Pro Ala Leu Ala Ser Leu Ser Arg Ser Ser Ser	
85 90 95	
ggc agg tca tca tcc gcc agg tca gcc tcg gtg aca acc tcc cca acc	336
Gly Arg Ser Ser Ser Ala Arg Ser Ala Ser Val Thr Thr Ser Pro Thr	
100 105 110	
aga gtg tac ctt gtt aga gca aca cca gtg ggg gct gta ccc atc cga	384
Arg Val Tyr Leu Val Arg Ala Thr Pro Val Gly Ala Val Pro Ile Arg	
115 120 125	
tca tct cct gcc agg tca gca cca gca acc agg gcc acc agg gag agc	432
Ser Ser Pro Ala Arg Ser Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser	
130 135 140	
cca ggt acg agc ctg ccc aag ttc acc tgg cgg gag ggc cag aag cag	480
Pro Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln	
145 150 155 160	
cta ccg ctc atc ggg tgc gtg ctc ctc ctc att gcc ctg gtg gtt tcg	528
Leu Pro Leu Ile Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser	
165 170 175	
ctc atc atc ctc ttc cag ttc tgg cag ggc cac aca ggg atc agg tac	576
Leu Ile Ile Leu Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr	
180 185 190	

-43-

aag gag cag agg gag agc tgt ccc aag cac gct gtt cgc tgt gac ggg Lys Glu Gln Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly	624
195 200 205	
gtg gtg gac tgc aag ctg aag agt gac gag ctg ggc tgc gtg agg ttt Val Val Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe	672
210 215 220	
gac tgg gac aag tct ctg ctt aaa atc tac tct ggg tcc tcc cat cag Asp Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His Gln	720
225 230 235 240	
tgg ctt ccc atc tgt agc agc aac tgg aat gac tcc tac tca gag aag Trp Leu Pro Ile Cys Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu Lys	768
245 250 255	
acc tgc cag cag ctg ggt ttc gag agt gct cac cgg aca acc gag gtt Thr Cys Gln Gln Leu Gly Phe Glu Ser Ala His Arg Thr Thr Glu Val	816
260 265 270	
gcc cac agg gat ttt gcc aac agc ttc tca atc ttg aga tac aac tcc Ala His Arg Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg Tyr Asn Ser	864
275 280 285	
acc atc cag gaa agc ctc cac agg tct gaa tgc cct tcc cag cgg tat Thr Ile Gln Glu Ser Leu His Arg Ser Glu Cys Pro Ser Gln Arg Tyr	912
290 295 300	
atc tcc ctc cag tgt tcc cac tgc gga ctg agg gcc atg acc ggg cgg Ile Ser Leu Gln Cys Ser His Cys Gly Leu Arg Ala Met Thr Gly Arg	960
305 310 315 320	
atc gtg gga ggg gcg ctg gcc tcg gat agc aag tgg cct tgg caa gtg Ile Val Gly Gly Ala Leu Ala Ser Asp Ser Lys Trp Pro Trp Gln Val	1008
325 330 335	
agt ctg cac ttc ggc acc acc cac atc tgt gga ggc acg ctc att gac Ser Leu His Phe Gly Thr Thr His Ile Cys Gly Gly Thr Leu Ile Asp	1056
340 345 350	
gcc cag tgg gtg ctc act gcc gcc cac tgc ttc ttc gtg acc cgg gag Ala Gln Trp Val Leu Thr Ala Ala His Cys Phe Phe Val Thr Arg Glu	1104
355 360 365	
aag gtc ctg gag ggc tgg aag gtg tac gcg ggc acc agc aac ctg cac Lys Val Leu Glu Gly Trp Lys Val Tyr Ala Gly Thr Ser Asn Leu His	1152
370 375 380	
cag ttg cct gag gca gcc tcc att gcc gag atc atc atc aac agc aat Gln Leu Pro Glu Ala Ala Ser Ile Ala Glu Ile Ile Ile Asn Ser Asn	1200
385 390 395 400	
tac acc gat gag gag gac gac tat gac atc gcc ctc atg cgg ctg tcc Tyr Thr Asp Glu Glu Asp Asp Tyr Asp Ile Ala Leu Met Arg Leu Ser	1248
405 410 415	
aag ccc ctg acc ctg tcc gct cac atc cac cct gct tgc ctc ccc atg Lys Pro Leu Thr Leu Ser Ala His Ile His Pro Ala Cys Leu Pro Met	1296
420 425 430	

-44-

```

cat gga cag acc ttt agc ctc aat gag acc tgc tgg atc aca ggc ttt      1344
His Gly Gln Thr Phe Ser Leu Asn Glu Thr Cys Trp Ile Thr Gly Phe
      435              440              445

ggc aag acc agg gag aca gat gac aag aca tcc ccc ttc ctc cgg gag      1392
Gly Lys Thr Arg Glu Thr Asp Asp Lys Thr Ser Pro Phe Leu Arg Glu
      450              455              460

gtg cag gtc aat ctc atc gac ttc aag aaa tgc aat gac tac ttg gtc      1440
Val Gln Val Asn Leu Ile Asp Phe Lys Lys Cys Asn Asp Tyr Leu Val
      465              470              475              480

tat gac agt tac ctt acc cca agg atg atg tgt gct ggg gac ctt cgt      1488
Tyr Asp Ser Tyr Leu Thr Pro Arg Met Met Cys Ala Gly Asp Leu Arg
      485              490              495

ggg ggc aga gac tcc tgc cag gga gac agc ggg ggg cct ctt gtc tgt      1536
Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys
      500              505              510

gag cag aac aac cgc tgg tac ctg gca ggt gtc acc agc tgg ggc aca      1584
Glu Gln Asn Asn Arg Trp Tyr Leu Ala Gly Val Thr Ser Trp Gly Thr
      515              520              525

ggc tgt ggc cag aga aac aaa cct ggt gtg tac acc aaa gtg aca gaa      1632
Gly Cys Gly Gln Arg Asn Lys Pro Gly Val Tyr Thr Lys Val Thr Glu
      530              535              540

gtt ctt ccc tgg att tac agc aag atg gag agc gag gtg cga ttc ata      1680
Val Leu Pro Trp Ile Tyr Ser Lys Met Glu Ser Glu Val Arg Phe Ile
      545              550              555              560

aaa tcc taa
Lys Ser *

```

<210> 24  
 <211> 562  
 <212> PRT  
 <213> homo sapien

<220>  
 <221> protease domain of endotheliase 2  
 <222> (321)..(562)

```

<400> 4
Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro Ser
  1              5              10              15
Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro Pro Gly
      20              25              30
Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly
      35              40              45
Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr
      50              55              60
Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser
      65              70              75              80
Pro Ala Arg Ala Ser Pro Ala Leu Ala Ser Leu Ser Arg Ser Ser Ser
      85              90              95
Gly Arg Ser Ser Ala Arg Ser Ala Ser Val Thr Thr Ser Pro Thr
      100              105              110
Arg Val Tyr Leu Val Arg Ala Thr Pro Val Gly Ala Val Pro Ile Arg
      115              120              125

```

-45-

Ser Ser Pro Ala Arg Ser Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser  
 130 135 140  
 Pro Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln  
 145 150 155 160  
 Leu Pro Leu Ile Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser  
 165 170 175  
 Leu Ile Ile Leu Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr  
 180 185 190  
 Lys Glu Gln Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly  
 195 200 205  
 Val Val Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe  
 210 215 220  
 Asp Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His Gln  
 225 230 235 240  
 Trp Leu Pro Ile Cys Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu Lys  
 245 250 255  
 Thr Cys Gln Gln Leu Gly Phe Glu Ser Ala His Arg Thr Thr Glu Val  
 260 265 270  
 Ala His Arg Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg Tyr Asn Ser  
 275 280 285  
 Thr Ile Gln Glu Ser Leu His Arg Ser Glu Cys Pro Ser Gln Arg Tyr  
 290 295 300  
 Ile Ser Leu Gln Cys Ser His Cys Gly Leu Arg Ala Met Thr Gly Arg  
 305 310 315 320  
 Ile Val Gly Gly Ala Leu Ala Ser Asp Ser Lys Trp Pro Trp Gln Val  
 325 330 335  
 Ser Leu His Phe Gly Thr Thr His Ile Cys Gly Gly Thr Leu Ile Asp  
 340 345 350  
 Ala Gln Trp Val Leu Thr Ala Ala His Cys Phe Phe Val Thr Arg Glu  
 355 360 365  
 Lys Val Leu Glu Gly Trp Lys Val Tyr Ala Gly Thr Ser Asn Leu His  
 370 375 380  
 Gln Leu Pro Glu Ala Ala Ser Ile Ala Glu Ile Ile Ile Asn Ser Asn  
 385 390 395 400  
 Tyr Thr Asp Glu Glu Asp Asp Tyr Asp Ile Ala Leu Met Arg Leu Ser  
 405 410 415  
 Lys Pro Leu Thr Leu Ser Ala His Ile His Pro Ala Cys Leu Pro Met  
 420 425 430  
 His Gly Gln Thr Phe Ser Leu Asn Glu Thr Cys Trp Ile Thr Gly Phe  
 435 440 445  
 Gly Lys Thr Arg Glu Thr Asp Asp Lys Thr Ser Pro Phe Leu Arg Glu  
 450 455 460  
 Val Gln Val Asn Leu Ile Asp Phe Lys Lys Cys Asn Asp Tyr Leu Val  
 465 470 475 480  
 Tyr Asp Ser Tyr Leu Thr Pro Arg Met Met Cys Ala Gly Asp Leu Arg  
 485 490 495  
 Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys  
 500 505 510  
 Glu Gln Asn Asn Arg Trp Tyr Leu Ala Gly Val Thr Ser Trp Gly Thr  
 515 520 525  
 Gly Cys Gly Gln Arg Asn Lys Pro Gly Val Tyr Thr Lys Val Thr Glu  
 530 535 540  
 Val Leu Pro Trp Ile Tyr Ser Lys Met Glu Ser Glu Val Arg Phe Ile  
 545 550 555 560  
 Lys Ser

&lt;210&gt; 25

&lt;211&gt; 2067

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien



-46-

<220>  
 <221> CDS  
 <222> (1)...(2067)  
 <223> Nucleic acid encoding (endotheliase 2-L) protein  
  
 <400> 25  
 atg gag agg gac agc cac ggg aat gca tct cca gca aga aca cct tca 48  
 Met Glu Arg Asp Ser His Gly Asn Ala Ser Pro Ala Arg Thr Pro Ser  
 1 5 10 15  
  
 gct gga gca tct cca gcc cag gca tct cca gct ggg aca cct cca ggc 96  
 Ala Gly Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr Pro Pro Gly  
 20 25 30  
  
 cgg gca tct cca gcc cag gca tct cca gcc cag gca tct cca gct ggg 144  
 Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly  
 35 40 45  
  
 aca cct ccg ggc cgg gca tct cca gcc cag gca tct cca gct ggt aca 192  
 Thr Pro Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser Pro Ala Gly Thr  
 50 55 60  
  
 cct cca ggc cgg gca tct cca gcc cgg gca tct cca gcc cag gca tct 240  
 Pro Pro Gly Arg Ala Ser Pro Gly Arg Ala Ser Pro Ala Gln Ala Ser  
 65 70 75 80  
  
 cca gcc cgg gca tct ccg gct ctg gca tca ctt tcc agg tcc tca tcc 288  
 Pro Ala Arg Ala Ser Pro Ala Leu Ala Ser Leu Ser Arg Ser Ser Ser  
 85 90 95  
  
 ggc agg tca tca tcc gcc agg tca gcc tcg gtg aca acc tcc cca acc 336  
 Gly Arg Ser Ser Ser Ala Arg Ser Ala Ser Val Thr Thr Ser Pro Thr  
 100 105 110  
  
 aga gtg tac ctt gtt aga gca aca cca gtg ggg gct gta ccc atc cga 384  
 Arg Val Tyr Leu Val Arg Ala Thr Pro Val Gly Ala Val Pro Ile Arg  
 115 120 125  
  
 tca tct cct gcc agg tca gca cca gca acc agg gcc acc agg gag agc 432  
 Ser Ser Pro Ala Arg Ser Ala Pro Ala Thr Arg Ala Thr Arg Glu Ser  
 130 135 140  
  
 cca ggt acg agc ctg ccc aag ttc acc tgg cgg gag ggc cag aag cag 480  
 Pro Gly Thr Ser Leu Pro Lys Phe Thr Trp Arg Glu Gly Gln Lys Gln  
 145 150 155 160  
  
 cta ccg ctc atc ggg tgc gtg ctc ctc ctc att gcc ctg gtg gtt tcg 528  
 Leu Pro Leu Ile Gly Cys Val Leu Leu Leu Ile Ala Leu Val Val Ser  
 165 170 175  
  
 ctc atc atc ctc ttc cag ttc tgg cag ggc cac aca ggg atc agg tac 576  
 Leu Ile Ile Leu Phe Gln Phe Trp Gln Gly His Thr Gly Ile Arg Tyr  
 180 185 190  
  
 aag gag cag agg gag agc tgt ccc aag cac gct gtt cgc tgt gac ggg 624  
 Lys Glu Gln Arg Glu Ser Cys Pro Lys His Ala Val Arg Cys Asp Gly  
 195 200 205  
  
 gtg gtg gac tgc aag ctg aag agt gac gag ctg ggc tgc gtg agg ttt 672  
 Val Val Asp Cys Lys Leu Lys Ser Asp Glu Leu Gly Cys Val Arg Phe  
 210 215 220

-47-

gac tgg gac aag tct ctg ctt aaa atc tac tct ggg tcc tcc cat cag Asp Trp Asp Lys Ser Leu Leu Lys Ile Tyr Ser Gly Ser Ser His Gln 225 230 235 240	720
tgg ctt ccc atc tgt agc agc aac tgg aat gac tcc tac tca gag aag Trp Leu Pro Ile Cys Ser Ser Asn Trp Asn Asp Ser Tyr Ser Glu Lys 245 250 255	768
acc tgc cag cag ctg ggt ttc gag agt gct cac cgg aca acc gag gtt Thr Cys Gln Gln Leu Gly Phe Glu Ser Ala His Arg Thr Thr Glu Val 260 265 270	816
gcc cac agg gat ttt gcc aac agc ttc tca atc ttg aga tac aac tcc Ala His Arg Asp Phe Ala Asn Ser Phe Ser Ile Leu Arg Tyr Asn Ser 275 280 285	864
acc atc cag gaa agc ctc cac agg tct gaa tgc cct tcc cag cgg tat Thr Ile Gln Glu Ser Leu His Arg Ser Glu Cys Pro Ser Gln Arg Tyr 290 295 300	912
atc tcc ctc cag tgt tcc cac tgc gga ctg agg gcc atg acc ggg cgg Ile Ser Leu Gln Cys Ser His Cys Gly Leu Arg Ala Met Thr Gly Arg 305 310 315 320	960
atc gtg gga ggg gcg ctg gcc tcg gat agc aag tgg cct tgg caa gtg Ile Val Gly Gly Ala Leu Ala Ser Asp Ser Lys Trp Pro Trp Gln Val 325 330 335	1008
agt ctg cac ttc ggc acc acc cac atc tgt gga ggc acg ctc att gac Ser Leu His Phe Gly Thr Thr His Ile Cys Gly Gly Thr Leu Ile Asp 340 345 350	1056
gcc cag tgg gtg ctc act gcc gcc cac tgc ttc ttc gtg acc cgg gag Ala Gln Trp Val Leu Thr Ala Ala His Cys Phe Phe Val Thr Arg Glu 355 360 365	1104
aag gtc ctg gag ggc tgg aag gtg tac gcg ggc acc agc aac ctg cac Lys Val Leu Glu Gly Trp Lys Val Tyr Ala Gly Thr Ser Asn Leu His 370 375 380	1152
cag ttg cct gag gca gcc tcc att gcc gag atc atc atc aac agc aat Gln Leu Pro Glu Ala Ala Ser Ile Ala Glu Ile Ile Ile Asn Ser Asn 385 390 395 400	1200
tac acc gat gag gag gac gac tat gac atc gcc ctc atg cgg ctg tcc Tyr Thr Asp Glu Glu Asp Asp Tyr Asp Ile Ala Leu Met Arg Leu Ser 405 410 415	1248
aag ccc ctg acc ctg tcc gct cac atc cac cct gct tgc ctc ccc atg Lys Pro Leu Thr Leu Ser Ala His Ile His Pro Ala Cys Leu Pro Met 420 425 430	1296
cat gga cag acc ttt agc ctc aat gag acc tgc tgg atc aca ggc ttt His Gly Gln Thr Phe Ser Leu Asn Glu Thr Cys Trp Ile Thr Gly Phe 435 440 445	1344
ggc aag acc agg gag aca gat gac aag aca tcc ccc ttc ctc cgg gag Gly Lys Thr Arg Glu Thr Asp Asp Lys Thr Ser Pro Phe Leu Arg Glu 450 455 460	1392

gtg	cag	gtc	aat	ctc	atc	gac	ttc	aag	aaa	tgc	aat	gac	tac	ttg	gtc		1440
Val	Gln	Val	Asn	Leu	Ile	Asp	Phe	Lys	Lys	Cys	Asn	Asp	Tyr	Leu	Val		.
465					470					475					480		
tat	gac	agt	tac	ctt	acc	cca	agg	atg	atg	tgt	gct	ggg	gac	ctt	cgt		1488
Tyr	Asp	Ser	Tyr	Leu	Thr	Pro	Arg	Met	Met	Cys	Ala	Gly	Asp	Leu	Arg		
				485					490					495			
ggg	ggc	aga	gac	tcc	tgc	cag	gga	gac	agc	ggg	ggg	cct	ctt	gtc	tgt		1536
Gly	Gly	Arg	Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys		
			500					505					510				
gag	cag	aac	aac	cgc	tgg	tac	ctg	gca	ggt	gtc	acc	agc	tgg	ggc	aca		1584
Glu	Gln	Asn	Asn	Arg	Trp	Tyr	Leu	Ala	Gly	Val	Thr	Ser	Trp	Gly	Thr		
			515				520					525					
ggc	tgt	ggc	cag	aga	aac	aaa	cct	ggt	gtg	tac	acc	aaa	gtg	aca	gaa		1632
Gly	Cys	Gly	Gln	Arg	Asn	Lys	Pro	Gly	Val	Tyr	Thr	Lys	Val	Thr	Glu		
	530					535					540						
gtt	ctt	ccc	tgg	att	tac	agc	aag	atg	gag	aac	aga	gct	cag	cgg	gtt		1680
Val	Leu	Pro	Trp	Ile	Tyr	Ser	Lys	Met	Glu	Asn	Arg	Ala	Gln	Arg	Val		
545					550					555					560		
gaa	aaa	gcg	tgg	acc	tac	agg	cca	ggc	agg	cag	ttg	ctg	ggc	aga	tgt		1728
Glu	Lys	Ala	Trp	Thr	Tyr	Arg	Pro	Gly	Arg	Gln	Leu	Leu	Gly	Arg	Cys		
				565				570						575			
tct	ccc	aga	agt	att	ttt	ttg	tgt	aag	gtt	gca	atg	gac	ttt	gaa	aac		1776
Ser	Pro	Arg	Ser	Ile	Phe	Leu	Cys	Lys	Val	Ala	Met	Asp	Phe	Glu	Asn		
			580					585					590				
gtt	tca	gtt	tct	gca	gag	gat	ttt	gtg	ata	gtt	ttt	gtt	atc	aag	cat		1824
Val	Ser	Val	Ser	Ala	Glu	Asp	Phe	Val	Ile	Val	Phe	Val	Ile	Lys	His		
			595				600					605					
tta	tgc	atg	gga	atc	cgc	tct	tca	tgg	cct	ttc	cca	gct	ctg	ttt	gtt		1872
Leu	Cys	Met	Gly	Ile	Arg	Ser	Ser	Trp	Pro	Phe	Pro	Ala	Leu	Phe	Val		
	610					615					620						
tta	gtc	ttt	ttg	att	ttc	ttt	ttg	ttg	ttg	ttg	ttg	tct	ttt	tta	aaa		1920
Leu	Val	Phe	Leu	Ile	Phe	Phe	Leu	Leu	Leu	Leu	Leu	Ser	Phe	Leu	Lys		
625					630					635					640		
aac	aca	agt	gac	tcc	att	ttg	act	ctg	aca	act	ttc	aca	gct	gtc	acc		1968
Asn	Thr	Ser	Asp	Ser	Ile	Leu	Thr	Leu	Thr	Thr	Phe	Thr	Ala	Val	Thr		
				645					650					655			
aga	atg	ctc	cct	gag	aac	tac	cat	tct	ttc	cct	ttc	cca	ctt	aaa	ata		2016
Arg	Met	Leu	Pro	Glu	Asn	Tyr	His	Ser	Phe	Pro							

-49-

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; protease domain

&lt;222&gt; (321)..(688)

&lt;400&gt; 26

Met	Glu	Arg	Asp	Ser	His	Gly	Asn	Ala	Ser	Pro	Ala	Arg	Thr	Pro	Ser	1	5	10	15
Ala	Gly	Ala	Ser	Pro	Ala	Gln	Ala	Ser	Pro	Ala	Gly	Thr	Pro	Pro	Gly	20	25	30	
Arg	Ala	Ser	Pro	Ala	Gln	Ala	Ser	Pro	Ala	Gln	Ala	Ser	Pro	Ala	Gly	35	40	45	
Thr	Pro	Pro	Gly	Arg	Ala	Ser	Pro	Ala	Gln	Ala	Ser	Pro	Ala	Gly	Thr	50	55	60	
Pro	Pro	Gly	Arg	Ala	Ser	Pro	Gly	Arg	Ala	Ser	Pro	Ala	Gln	Ala	Ser	65	70	75	80
Pro	Ala	Arg	Ala	Ser	Pro	Ala	Leu	Ala	Ser	Leu	Ser	Arg	Ser	Ser	Ser	85	90	95	
Gly	Arg	Ser	Ser	Ser	Ala	Arg	Ser	Ala	Ser	Val	Thr	Thr	Ser	Pro	Thr	100	105	110	
Arg	Val	Tyr	Leu	Val	Arg	Ala	Thr	Pro	Val	Gly	Ala	Val	Pro	Ile	Arg	115	120	125	
Ser	Ser	Pro	Ala	Arg	Ser	Ala	Pro	Ala	Thr	Arg	Ala	Thr	Arg	Glu	Ser	130	135	140	
Pro	Gly	Thr	Ser	Leu	Pro	Lys	Phe	Thr	Trp	Arg	Glu	Gly	Gln	Lys	Gln	145	150	155	160
Leu	Pro	Leu	Ile	Gly	Cys	Val	Leu	Leu	Leu	Ile	Ala	Leu	Val	Val	Ser	165	170	175	
Leu	Ile	Ile	Leu	Phe	Gln	Phe	Trp	Gln	Gly	His	Thr	Gly	Ile	Arg	Tyr	180	185	190	
Lys	Glu	Gln	Arg	Glu	Ser	Cys	Pro	Lys	His	Ala	Val	Arg	Cys	Asp	Gly	195	200	205	
Val	Val	Asp	Cys	Lys	Leu	Lys	Ser	Asp	Glu	Leu	Gly	Cys	Val	Arg	Phe	210	215	220	
Asp	Trp	Asp	Lys	Ser	Leu	Leu	Lys	Ile	Tyr	Ser	Gly	Ser	Ser	His	Gln	225	230	235	240
Trp	Leu	Pro	Ile	Cys	Ser	Ser	Asn	Trp	Asn	Asp	Ser	Tyr	Ser	Glu	Lys	245	250	255	
Thr	Cys	Gln	Gln	Leu	Gly	Phe	Glu	Ser	Ala	His	Arg	Thr	Thr	Glu	Val	260	265	270	
Ala	His	Arg	Asp	Phe	Ala	Asn	Ser	Phe	Ser	Ile	Leu	Arg	Tyr	Asn	Ser	275	280	285	
Thr	Ile	Gln	Glu	Ser	Leu	His	Arg	Ser	Glu	Cys	Pro	Ser	Gln	Arg	Tyr	290	295	300	
Ile	Ser	Leu	Gln	Cys	Ser	His	Cys	Gly	Leu	Arg	Ala	Met	Thr	Gly	Arg	305	310	315	320
Ile	Val	Gly	Gly	Ala	Leu	Ala	Ser	Asp	Ser	Lys	Trp	Pro	Trp	Gln	Val	325	330	335	
Ser	Leu	His	Phe	Gly	Thr	Thr	His	Ile	Cys	Gly	Gly	Thr	Leu	Ile	Asp	340	345	350	
Ala	Gln	Trp	Val	Leu	Thr	Ala	Ala	His	Cys	Phe	Phe	Val	Thr	Arg	Glu	355	360	365	
Lys	Val	Leu	Glu	Gly	Trp	Lys	Val	Tyr	Ala	Gly	Thr	Ser	Asn	Leu	His	370	375	380	
Gln	Leu	Pro	Glu	Ala	Ala	Ser	Ile	Ala	Glu	Ile	Ile	Ile	Asn	Ser	Asn	385	390	395	400
Tyr	Thr	Asp	Glu	Glu	Asp	Asp	Tyr	Asp	Ile	Ala	Leu	Met	Arg	Leu	Ser	405	410	415	
Lys	Pro	Leu	Thr	Leu	Ser	Ala	His	Ile	His	Pro	Ala	Cys	Leu	Pro	Met	420	425	430	

-50-

His Gly Gln Thr Phe Ser Leu Asn Glu Thr Cys Trp Ile Thr Gly Phe  
 435 440 445  
 Gly Lys Thr Arg Glu Thr Asp Asp Lys Thr Ser Pro Phe Leu Arg Glu  
 450 455 460  
 Val Gln Val Asn Leu Ile Asp Phe Lys Lys Cys Asn Asp Tyr Leu Val  
 465 470 475 480  
 Tyr Asp Ser Tyr Leu Thr Pro Arg Met Met Cys Ala Gly Asp Leu Arg  
 485 490 495  
 Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys  
 500 505 510  
 Glu Gln Asn Asn Arg Trp Tyr Leu Ala Gly Val Thr Ser Trp Gly Thr  
 515 520 525  
 Gly Cys Gly Gln Arg Asn Lys Pro Gly Val Tyr Thr Lys Val Thr Glu  
 530 535 540  
 Val Leu Pro Trp Ile Tyr Ser Lys Met Glu Asn Arg Ala Gln Arg Val  
 545 550 555 560  
 Glu Lys Ala Trp Thr Tyr Arg Pro Gly Arg Gln Leu Leu Gly Arg Cys  
 565 570 575  
 Ser Pro Arg Ser Ile Phe Leu Cys Lys Val Ala Met Asp Phe Glu Asn  
 580 585 590  
 Val Ser Val Ser Ala Glu Asp Phe Val Ile Val Phe Val Ile Lys His  
 595 600 605  
 Leu Cys Met Gly Ile Arg Ser Ser Trp Pro Phe Pro Ala Leu Phe Val  
 610 615 620  
 Leu Val Phe Leu Ile Phe Phe Leu Leu Leu Leu Ser Phe Leu Lys  
 625 630 635 640  
 Asn Thr Ser Asp Ser Ile Leu Thr Leu Thr Thr Phe Thr Ala Val Thr  
 645 650 655  
 Arg Met Leu Pro Glu Asn Tyr His Ser Phe Pro Phe Pro Leu Lys Ile  
 660 665 670  
 Phe His Gln Asn Leu Thr Thr Ile Ile Lys Glu Tyr Lys Val Ile Lys  
 675 680 685

<210> 27  
 <211> 1471  
 <212> DNA  
 <213> Homo Sapien

<220>  
 <223> DESC1 gene  
 <221> misc\_feature  
 <222> (626)...(1324)  
 <223> protease domain

<221> CDS  
 <222> (56)...(1324)

<400> 27  
 tgacttgat gtagacctcg accttcacag gactcttcat tgctgggttg caatg atg 58  
 Met  
 1  
 tat cgg cca gat gtg gtg agg gct agg aaa aga gtt tgt tgg gaa ccc 106  
 Tyr Arg Pro Asp Val Val Arg Ala Arg Lys Arg Val Cys Trp Glu Pro  
 5 10 15  
 tgg gtt atc ggc ctc gtc ats ttc ata tcc ctg att gtc ctg gca gtg 154  
 Trp Val Ile Gly Leu Val Xaa Phe Ile Ser Leu Ile Val Leu Ala Val  
 20 25 30

-51-

tgc att gga stc act gtt cat tat gtg aga tat aat caa aag aag acc Cys Ile Gly Xaa Thr Val His Tyr Val Arg Tyr Asn Gln Lys Lys Thr 35 40 45	202
tac aat tac tat agc aca ttg tca ttt aca act gac aaa cta tat gct Tyr Asn Tyr Tyr Ser Thr Leu Ser Phe Thr Thr Asp Lys Leu Tyr Ala 50 55 60 65	250
gag ttt ggc aga gag gct tct aac aat ttt aca gaa atg agc cag aga Glu Phe Gly Arg Glu Ala Ser Asn Asn Phe Thr Glu Met Ser Gln Arg 70 75 80	298
ctt gaa tca atg gtg aaa aat gca ttt tat aaa tct cca tta agg gaa Leu Glu Ser Met Val Lys Asn Ala Phe Tyr Lys Ser Pro Leu Arg Glu 85 90 95	346
gaa ttt gtc aag tct cag gtt atc aag ttc agt caa cag aag cat gga Glu Phe Val Lys Ser Gln Val Ile Lys Phe Ser Gln Gln Lys His Gly 100 105 110	394
gtg ttg gct cat atg ctg ttg att tgt aga ttt cac tct act gag gat Val Leu Ala His Met Leu Leu Ile Cys Arg Phe His Ser Thr Glu Asp 115 120 125	442
cct gaa act gta gat aaa att gtt caa ctt gtt tta cat gaa aag ctg Pro Glu Thr Val Asp Lys Ile Val Gln Leu Val Leu His Glu Lys Leu 130 135 140 145	490
caa gat gct gta gga ccc cct aaa gta gat cct cac tca gtt aaa att Gln Asp Ala Val Gly Pro Pro Lys Val Asp Pro His Ser Val Lys Ile 150 155 160	538
aaa aaa atc aac aag aca gaa aca gac agc tat cta aac cat tgc tgc Lys Lys Ile Asn Lys Thr Glu Thr Asp Ser Tyr Leu Asn His Cys Cys 165 170 175	586
gga aca cga aga agt aaa act cta ggt cag agt ctc agg atc gtt ggt Gly Thr Arg Arg Ser Lys Thr Leu Gly Gln Ser Leu Arg Ile Val Gly 180 185 190	634
ggg aca gaa gta gaa gag ggt gaa tgg ccc tgg cag gct agc ctg cag Gly Thr Glu Val Glu Glu Gly Glu Trp Pro Trp Gln Ala Ser Leu Gln 195 200 205	682
tgg gat ggg agt cat cgc tgt gga gca acc tta att aat gcc aca tgg Trp Asp Gly Ser His Arg Cys Gly Ala Thr Leu Ile Asn Ala Thr Trp 210 215 220 225	730
ctt gtg agt gct gct cac tgt ttt aca aca tat aag aac cct gcc aga Leu Val Ser Ala Ala His Cys Phe Thr Thr Tyr Lys Asn Pro Ala Arg 230 235 240	778
tgg act gct tcc ttt gga gta aca ata aaa cct tcg aaa atg aaa cgg Trp Thr Ala Ser Phe Gly Val Thr Ile Lys Pro Ser Lys Met Lys Arg 245 250 255	826
ggt ctc cgg aga ata att gtc cat gaa aaa tac aaa cac cca tca cat Gly Leu Arg Arg Ile Ile Val His Glu Lys Tyr Lys His Pro Ser His 260 265 270	874
gac tat gat att tct ctt gca gag ctt tct agc cct gtt ccc tac aca	922

[illegible]

-53-

gct ccg gaa gag cgc tac cgc aga gcc ggg tcc cca aag ccg gtc ttg	162
Ala Pro Glu Glu Arg Tyr Arg Arg Ala Gly Ser Pro Lys Pro Val Leu	
10 15 20	
aga gct gat gac aat aac atg ggc aat ggc tgc tct cag aag ctg gcg	210
Arg Ala Asp Asp Asn Asn Met Gly Asn Gly Cys Ser Gln Lys Leu Ala	
25 30 35	
act gct aac ctc ctc cgg ttc cta ttg ctg gtc ctg att cca tgt atc	258
Thr Ala Asn Leu Leu Arg Phe Leu Leu Leu Val Leu Ile Pro Cys Ile	
40 45 50 55	
tgt gct ctc gtt ctc ttg ctg gtg atc ctg ctt tcc tat gtt gga aca	306
Cys Ala Leu Val Leu Leu Leu Val Ile Leu Leu Ser Tyr Val Gly Thr	
60 65 70	
tta caa aag gtc tat ttt aaa tca aat ggg agt gaa cct ttg gtc act	354
Leu Gln Lys Val Tyr Phe Lys Ser Asn Gly Ser Glu Pro Leu Val Thr	
75 80 85	
gat ggt gaa atc caa ggg tcc gat gtt att ctt aca aat aca att tat	402
Asp Gly Glu Ile Gln Gly Ser Asp Val Ile Leu Thr Asn Thr Ile Tyr	
90 95 100	
aac cag agc act gtg gtg tct act gca cat ccc gac caa cac gtt cca	450
Asn Gln Ser Thr Val Val Ser Thr Ala His Pro Asp Gln His Val Pro	
105 110 115	
gcc tgg act acg gat gct tct ctc cca ggg gac caa agt cac agg aat	498
Ala Trp Thr Thr Asp Ala Ser Leu Pro Gly Asp Gln Ser His Arg Asn	
120 125 130 135	
aca agt gcc tgt atg aac atc acc cac agc cag tgt cag atg ctg ccc	546
Thr Ser Ala Cys Met Asn Ile Thr His Ser Gln Cys Gln Met Leu Pro	
140 145 150	
tac cac gcc acg ctg aca cct ctc ctc tca gtt gtc aga aac atg gaa	594
Tyr His Ala Thr Leu Thr Pro Leu Leu Ser Val Val Arg Asn Met Glu	
155 160 165	
atg gaa aag ttc ctc aag ttt ttc aca tat ctc cat cgc ctc agt tgc	642
Met Glu Lys Phe Leu Lys Phe Phe Thr Tyr Leu His Arg Leu Ser Cys	
170 175 180	
tat caa cat atc atg ctg ttt ggc tgt acc ctc gcc ttc cct gag tgc	690
Tyr Gln His Ile Met Leu Phe Gly Cys Thr Leu Ala Phe Pro Glu Cys	
185 190 195	
atc att gat ggc gat gac agt cat gga ctc ctg ccc tgt agg tcc ttc	738
Ile Ile Asp Gly Asp Asp Ser His Gly Leu Leu Pro Cys Arg Ser Phe	
200 205 210 215	
tgt gag gct gca aaa gaa ggc tgt gaa tca gtc ctg ggg atg gtg aat	786
Cys Glu Ala Ala Lys Glu Gly Cys Glu Ser Val Leu Gly Met Val Asn	
220 225 230	
tac tcc tgg ccg gat ttc ctc aga tgc tcc cag ttt aga aac caa act	834
Tyr Ser Trp Pro Asp Phe Leu Arg Cys Ser Gln Phe Arg Asn Gln Thr	
235 240 245	



-54-

gaa agc agc aat gtc agc aga att tgc ttc tca cct cag cag gaa aac	882
Glu Ser Ser Asn Val Ser Arg Ile Cys Phe Ser Pro Gln Gln Glu Asn	
250 255 260	
gga aag caa ttg ctc tgt gga agg ggt gag aac ttt ctg tgt gcc agt	930
Gly Lys Gln Leu Leu Cys Gly Arg Gly Glu Asn Phe Leu Cys Ala Ser	
265 270 275	
gga atc tgc atc ccc ggg aaa ctg caa tgt aat ggc tac aac gac tgt	978
Gly Ile Cys Ile Pro Gly Lys Leu Gln Cys Asn Gly Tyr Asn Asp Cys	
280 285 290 295	
gac gac tgg agt gac gag gct cat tgc aac tgc agc gag aat ctg ttt	1026
Asp Asp Trp Ser Asp Glu Ala His Cys Asn Cys Ser Glu Asn Leu Phe	
300 305 310	
cac tgt cac aca ggc aag tgc ctt aat tac agc ctt gtg tgt gat gga	1074
His Cys His Thr Gly Lys Cys Leu Asn Tyr Ser Leu Val Cys Asp Gly	
315 320 325	
tat gat gac tgt ggg gat ttg agt gat gag caa aac tgt gat tgc aat	1122
Tyr Asp Asp Cys Gly Asp Leu Ser Asp Glu Gln Asn Cys Asp Cys Asn	
330 335 340	
ccc aca aca gag cat cgc tgc ggg gac ggg cgc tgc atc gcc atg gag	1170
Pro Thr Thr Glu His Arg Cys Gly Asp Gly Arg Cys Ile Ala Met Glu	
345 350 355	
tgg gtg tgt gat ggt gac cac gac tgt gtg gat aag tcc gac gag gtc	1218
Trp Val Cys Asp Gly Asp His Asp Cys Val Asp Lys Ser Asp Glu Val	
360 365 370 375	
aac tgc tcc tgt cac agc cag ggt ctg gtg gaa tgc aga aat gga caa	1266
Asn Cys Ser Cys His Ser Gln Gly Leu Val Glu Cys Arg Asn Gly Gln	
380 385 390	
tgt atc ccc agc acg ttt caa tgt gat ggt gac gag gac tgc aag gat	1314
Cys Ile Pro Ser Thr Phe Gln Cys Asp Gly Asp Glu Asp Cys Lys Asp	
395 400 405	
ggg agt gat gag gag aac tgc agc gtc att cag act tca tgt caa gaa	1362
Gly Ser Asp Glu Glu Asn Cys Ser Val Ile Gln Thr Ser Cys Gln Glu	
410 415 420	
gga gac caa aga tgc ctc tac aat ccc tgc ctt gat tca tgt ggt ggt	1410
Gly Asp Gln Arg Cys Leu Tyr Asn Pro Cys Leu Asp Ser Cys Gly Gly	
425 430 435	
agc tct ctc tgt gac ccg aac aac agt ctg aat aac tgt agt caa tgt	1458
Ser Ser Leu Cys Asp Pro Asn Asn Ser Leu Asn Asn Cys Ser Gln Cys	
440 445 450 455	
gaa cca att aca ttg gaa ctc tgc atg aat ttg ccc tac aac agt aca	1506
Glu Pro Ile Thr Leu Glu Leu Cys Met Asn Leu Pro Tyr Asn Ser Thr	
460 465 470	
agt tat cca aat tat ttt ggc cac agg act caa aag gaa gca tcc atc	1554
Ser Tyr Pro Asn Tyr Phe Gly His Arg Thr Gln Lys Glu Ala Ser Ile	
475 480 485	
agc tgg gag tct tct ctt ttc cct gca ctt gtt caa acc aac tgt tat	1602

-55-

Ser	Trp	Glu	Ser	Ser	Leu	Phe	Pro	Ala	Leu	Val	Gln	Thr	Asn	Cys	Tyr	
		490					495					500				
aaa	tac	ctc	atg	ttc	ttt	tct	tgc	acc	att	ttg	gta	cca	aaa	tgt	gat	1650
Lys	Tyr	Leu	Met	Phe	Phe	Ser	Cys	Thr	Ile	Leu	Val	Pro	Lys	Cys	Asp	
	505					510					515					
gtg	aat	aca	ggc	gag	cgt	atc	cct	cct	tgc	agg	gca	ttg	tgt	gaa	cac	1698
Val	Asn	Thr	Gly	Glu	Arg	Ile	Pro	Pro	Cys	Arg	Ala	Leu	Cys	Glu	His	
	520				525					530					535	
tct	aaa	gaa	cgc	tgt	gag	tct	gtt	ctt	ggg	att	gtg	ggc	cta	cag	tgg	1746
Ser	Lys	Glu	Arg		Glu	Ser	Val	Leu	Gly	Ile	Val	Gly	Leu	Gln	Trp	
				540					545					550		
cct	gaa	gac	aca	gat	tgc	agt	caa	ttt	cca	gag	gaa	aat	tca	gac	aat	1794
Pro	Glu	Asp	Thr	Asp	Cys	Ser	Gln	Phe	Pro	Glu	Glu	Asn	Ser	Asp	Asn	
			555					560					565			
caa	acc	tgc	ctg	atg	cct	gat	gaa	tat	gtg	gaa	gaa	tgc	tca	cct	agt	1842
Gln	Thr	Cys	Leu	Met	Pro	Asp	Glu	Tyr	Val	Glu	Glu	Cys	Ser	Pro	Ser	
		570					575					580				
cat	ttc	aag	tgc	cgc	tca	gga	cag	tgt	gtt	ctg	gct	tcc	aga	aga	tgt	1890
His	Phe	Lys	Cys	Arg	Ser	Gly	Gln	Cys	Val	Leu	Ala	Ser	Arg	Arg	Cys	
	585					590					595					
gat	ggc	cag	gcc	gac	tgt	gac	gat	gac	agt	gat	gag	gaa	aac	tgt	ggc	1938
Asp	Gly	Gln	Ala	Asp	Cys	Asp	Asp	Asp	Ser	Asp	Glu	Glu	Asn	Cys	Gly	
	600				605					610					615	
tgt	aaa	gag	aga	gat	ctt	tgg	gaa	tgt	cca	tcc	aat	aaa	caa	tgt	ttg	1986
Cys	Lys	Glu	Arg	Asp	Leu	Trp	Glu	Cys	Pro	Ser	Asn	Lys	Gln	Cys	Leu	
				620					625					630		
aag	cac	aca	gtg	atc	tgc	gat	ggg	ttc	cca	gac	tgc	cct	gat	tac	atg	2034
Lys	His	Thr	Val	Ile	Cys	Asp	Gly	Phe	Pro	Asp	Cys	Pro	Asp	Tyr	Met	
			635					640					645			
gac	gag	aaa	aac	tgc	tca	ttt	tgc	caa	gat	gat	gag	ctg	gaa	tgt	gca	2082
Asp	Glu	Lys	Asn	Cys	Ser	Phe	Cys	Gln	Asp	Asp	Glu	Leu	Glu	Cys	Ala	
		650					655					660				
aac	cat	gcg	tgt	gtg	tca	cgt	gac	ctg	tgg	tgt	gat	ggc	gaa	gcc	gac	2130
Asn	His	Ala	Cys	Val	Ser	Arg	Asp	Leu	Trp	Cys	Asp	Gly	Glu	Ala	Asp	
	665					670					675					
tgc	tca	gac	agt	tca	gat	gaa	tgg	gac	tgt	gtg	acc	ctc	tct	ata	aat	2178
Cys	Ser	Asp	Ser	Ser	Asp	Glu	Trp	Asp	Cys	Val	Thr	Leu	Ser	Ile	Asn	
					685					690					695	
gtg	aac	tcc	tct	tcc	ttt	ctg	atg	gtt	cac	aga	gct	gcc	aca	gaa	cac	2226
Val	Asn	Ser	Ser	Ser	Phe	Leu	Met	Val	His	Arg	Ala	Ala	Thr	Glu	His	
					700				705					710		
cat	gtg	tgt	gca	gat	ggc	tgg	cag	gag	ata	ttg	agt	cag	ctg	gcc	tgc	2274
His	Val	Cys	Ala	Asp	Gly	Trp	Gln	Glu	Ile	Leu	Ser	Gln	Leu	Ala	Cys	
			715				720						725			
aag	cag	atg	ggc	tta	gga	gaa	cca	tct	gtg	acc	aaa	ttg	ata	cag	gaa	2322
Lys	Gln	Met	Gly	Leu	Gly	Glu	Pro	Ser	Val	Thr	Lys	Leu	Ile	Gln	Glu	

-56-

730	735	740	
cag gag aaa gag ccg cgg tgg ctg aca tta cac tcc aac tgg gag agc Gln Glu Lys Glu Pro Arg Trp Leu Thr Leu His Ser Asn Trp Glu Ser 745 750 755			2370
ctc aat ggg acc act tta cat gaa ctt cta gta aat ggg cag tct tgt Leu Asn Gly Thr Thr Leu His Glu Leu Leu Val Asn Gly Gln Ser Cys 760 765 770 775			2418
gag agc aga agt aaa att tct ctt ctg tgt act aaa caa gac tgt ggg Glu Ser Arg Ser Lys Ile Ser Leu Leu Cys Thr Lys Gln Asp Cys Gly 780 785 790			2466
cgc cgc cct gct gcc cga atg aac aaa agg atc ctt gga ggt cgg acg Arg Arg Pro Ala Ala Arg Met Asn Lys Arg Ile Leu Gly Gly Arg Thr 795 800 805			2514
agt cgc cct gga agg tgg cca tgg cag tgt tct ctg cag agt gaa ccc Ser Arg Pro Gly Arg Trp Pro Trp Gln Cys Ser Leu Gln Ser Glu Pro 810 815 820			2562
agt gga cat atc tgt ggc tgt gtc ctc att gcc aag aag tgg gtt ctg Ser Gly His Ile Cys Gly Cys Val Leu Ile Ala Lys Lys Trp Val Leu 825 830 835			2610
aca gtt gcc cac tgc ttc gag ggg aga gag aat gct gca gtt tgg aaa Thr Val Ala His Cys Phe Glu Gly Arg Glu Asn Ala Ala Val Trp Lys 840 845 850 855			2658
gtg gtg ctt ggc atc aac aat cta gac cat cca tca gtg ttc atg cag Val Val Leu Gly Ile Asn Asn Leu Asp His Pro Ser Val Phe Met Gln 860 865 870			2706
aca cgc ttt gtg aag acc atc atc ctg cat ccc cgc tac agt cga gca Thr Arg Phe Val Lys Thr Ile Ile Leu His Pro Arg Tyr Ser Arg Ala 875 880 885			2754
gtg gtg gac tat gac atc agc atc gtt gag ctg agt gaa gac atc agt Val Val Asp Tyr Asp Ile Ser Ile Val Glu Leu Ser Glu Asp Ile Ser 890 895 900			2802
gag act ggc tac gtc cgg cct gtc tgc ttg ccc aac ccg gag cag tgg Glu Thr Gly Tyr Val Arg Pro Val Cys Leu Pro Asn Pro Glu Gln Trp 905 910 915			2850
cta gag cct gac acg tac tgc tat atc aca ggc tgg ggc cac atg ggc Leu Glu Pro Asp Thr Tyr Cys Tyr Ile Thr Gly Trp Gly His Met Gly 920 925 930 935			2898
aat aaa atg cca ttt aag ctg caa gag gga gag gtc cgc att att tct Asn Lys Met Pro Phe Lys Leu Gln Glu Gly Glu Val Arg Ile Ile Ser 940 945 950			2946
ctg gaa cat tgt cag tcc tac ttt gac atg aag acc atc acc act cgg Leu Glu His Cys Gln Ser Tyr Phe Asp Met Lys Thr Ile Thr Thr Arg 955 960 965			2994
atg ata tgt gct ggc tat gag tct ggc aca gtt gat tca tgc atg ggt Met Ile Cys Ala Gly Tyr Glu Ser Gly Thr Val Asp Ser Cys Met Gly 970 975 980			3042

-57-

gac agc ggt ggg cct ctt gtt tgt gag aag cct gga gga cgg tgg aca 3090  
 Asp Ser Gly Gly Pro Leu Val Cys Glu Lys Pro Gly Gly Arg Trp Thr  
 985 990 995  
 tta ttt gga tta act tca tgg ggc tcc gtc tgc ttt tcc aaa gtc ctg 3138  
 Leu Phe Gly Leu Thr Ser Trp Gly Ser Val Cys Phe Ser Lys Val Leu  
 1000 1005 1010 1015  
 ggg cct ggc gtt tat agt aat gtg tca tat ttc gtc gaa tgg att aaa 3186  
 Gly Pro Gly Val Tyr Ser Asn Val Ser Tyr Phe Val Glu Trp Ile Lys  
 1020 1025 1030  
 aga cag att tac atc cag acc ttt ctc cta aac taa ttataaggat 3232  
 Arg Gln Ile Tyr Ile Gln Thr Phe Leu Leu Asn \*  
 1035 1040  
 gatcagagac ttttgccagc tacactaaaa gaaaatggcc ttcttgactg tgaagagctg 3292  
 cctgcagaga gctgtacaga agcacttttc atggacagaa atgctcaatc gtgcactgca 3352  
 aatttgcatg tttgttttgg actaattttt ttcaatttat tttttcacct tcatttttct 3412  
 cttattttcaa gttcaatgaa agactttaca aaagcaaaca aagcagactt tgccttttg 3472  
 ccaggcctaa ccatgactgc agcacaaaat tatcgactct ggcgagattt aaaatcaggt 3532  
 gctacagtaa cagggttatgg aatgggtctct tttatcctat cacaaaaaaa gacatagata 3592  
 tttaggctga ttaattatct ctaccagttt ttgttttctca agctcagtgc atagtggtaa 3652  
 atttcagtggt taacattgga gacttgcttt tctttttctt tttttatacc ccacaattct 3712  
 tttttattac acttcgaatt ttaggggtaca cgagcacaaac gtgcagggtta gttacatatg 3772  
 tatacatgtg ccatgttggt gtgctgaacc cagtaactcg tcatttgatt tattaaaagc 3832  
 caagataatt tacatgttta aagtatttac tattaccccc ttctaattgtt tgcataattc 3892  
 tgagaactga taaaagacag caataaaaaga ccagtgtcat ccatttaggt agcaagacat 3952  
 attgaatgca aagttcttta gatatacaata ttaacacttg acattattgg acccccatt 4012  
 ctggatgtat atcaagatca taattttata gaagagtctc tatagaactg tcctcatagc 4072  
 tgggtttgtt caggatatat gagttggctg attgagactg caacaactac atctatatat 4132  
 atgggcaata ttttggttta cttatgtggc aaagaactgg atattaaact ttgcaaaaaga 4192  
 gaatttagat gagagatgca atttttttaa aagaaaatta atttgcattc ctggtttaat 4252  
 taaattttatt tttcagtttt cttgcgttca tccataccaa caaagtcata aagagcatat 4312  
 ttttagagcac agtaagactt tgcattggagt aaaacatttt gtaattttcc tcaaaagatg 4372  
 tttaatatct ggtttcttct cattggtaat taaaatttta gaaatgattt ttagctctag 4432  
 gccactttac gcaactcaat ttctgaagca attagtggta aaaagtattt ttcccacta 4492  
 aaaaacttta aaacacaaat cttcatatat acttaattta attagtcagg catccatttt 4552  
 gcctttttaa caactaggat tccctactaa cctccaccag caacctggac tgcctcagca 4612  
 ttccaaatag atactacctg caattttata catgtatttt tgtatctttt ctgtgtgtaa 4672  
 acatagttga aattcaaaaa gttgtagcaa tttctatact attcatctcc tgtccttcag 4732  
 tttgtataaa cctaaggaga gtgtgaaatc cagcaactga attgtgggtca cgattgtatg 4792  
 aaagttcaag aacatatgtc agttttgtta cagttgtagc tacatactca atgtatcaac 4852  
 ttttagcctg ctcaacttag gctcagtga atatatatat tatacttatt ttaaataatt 4912  
 cttaatacaa ataaaatgggt a 4933

&lt;210&gt; 29

&lt;211&gt; 1042

&lt;212&gt; PRT

&lt;213&gt; Homo Sapien

&lt;400&gt; 29

Met Lys Gln Ser Pro Ala Leu Ala Pro Glu Glu Arg Tyr Arg Arg Ala  
 1 5 10 15  
 Gly Ser Pro Lys Pro Val Leu Arg Ala Asp Asp Asn Asn Met Gly Asn  
 20 25 30  
 Gly Cys Ser Gln Lys Leu Ala Thr Ala Asn Leu Leu Arg Phe Leu Leu  
 35 40 45  
 Leu Val Leu Ile Pro Cys Ile Cys Ala Leu Val Leu Leu Leu Val Ile  
 50 55 60

-58-

Leu	Leu	Ser	Tyr	Val	Gly	Thr	Leu	Gln	Lys	Val	Tyr	Phe	Lys	Ser	Asn
65					70					75					80
Gly	Ser	Glu	Pro	Leu	Val	Thr	Asp	Gly	Glu	Ile	Gln	Gly	Ser	Asp	Val
				85				90						95	
Ile	Leu	Thr	Asn	Thr	Ile	Tyr	Asn	Gln	Ser	Thr	Val	Val	Ser	Thr	Ala
			100					105					110		
His	Pro	Asp	Gln	His	Val	Pro	Ala	Trp	Thr	Thr	Asp	Ala	Ser	Leu	Pro
		115					120					125			
Gly	Asp	Gln	Ser	His	Arg	Asn	Thr	Ser	Ala	Cys	Met	Asn	Ile	Thr	His
	130					135					140				
Ser	Gln	Cys	Gln	Met	Leu	Pro	Tyr	His	Ala	Thr	Leu	Thr	Pro	Leu	Leu
145					150					155					160
Ser	Val	Val	Arg	Asn	Met	Glu	Met	Glu	Lys	Phe	Leu	Lys	Phe	Phe	Thr
				165					170					175	
Tyr	Leu	His	Arg	Leu	Ser	Cys	Tyr	Gln	His	Ile	Met	Leu	Phe	Gly	Cys
			180					185					190		
Thr	Leu	Ala	Phe	Pro	Glu	Cys	Ile	Ile	Asp	Gly	Asp	Asp	Ser	His	Gly
		195					200					205			
Leu	Leu	Pro	Cys	Arg	Ser	Phe	Cys	Glu	Ala	Ala	Lys	Glu	Gly	Cys	Glu
	210					215					220				
Ser	Val	Leu	Gly	Met	Val	Asn	Tyr	Ser	Trp	Pro	Asp	Phe	Leu	Arg	Cys
225					230					235					240
Ser	Gln	Phe	Arg	Asn	Gln	Thr	Glu	Ser	Ser	Asn	Val	Ser	Arg	Ile	Cys
				245					250					255	
Phe	Ser	Pro	Gln	Gln	Glu	Asn	Gly	Lys	Gln	Leu	Leu	Cys	Gly	Arg	Gly
			260				265						270		
Glu	Asn	Phe	Leu	Cys	Ala	Ser	Gly	Ile	Cys	Ile	Pro	Gly	Lys	Leu	Gln
		275					280					285			
Cys	Asn	Gly	Tyr	Asn	Asp	Cys	Asp	Asp	Trp	Ser	Asp	Glu	Ala	His	Cys
	290					295					300				
Asn	Cys	Ser	Glu	Asn	Leu	Phe	His	Cys	His	Thr	Gly	Lys	Cys	Leu	Asn
305					310					315					320
Tyr	Ser	Leu	Val	Cys	Asp	Gly	Tyr	Asp	Asp	Cys	Gly	Asp	Leu	Ser	Asp
				325					330					335	
Glu	Gln	Asn	Cys	Asp	Cys	Asn	Pro	Thr	Thr	Glu	His	Arg	Cys	Gly	Asp
			340				345						350		
Gly	Arg	Cys	Ile	Ala	Met	Glu	Trp	Val	Cys	Asp	Gly	Asp	His	Asp	Cys
		355					360					365			
Val	Asp	Lys	Ser	Asp	Glu	Val	Asn	Cys	Ser	Cys	His	Ser	Gln	Gly	Leu
	370					375					380				
Val	Glu	Cys	Arg	Asn	Gly	Gln	Cys	Ile	Pro	Ser	Thr	Phe	Gln	Cys	Asp
385					390					395					400
Gly	Asp	Glu	Asp	Cys	Lys	Asp	Gly	Ser	Asp	Glu	Glu	Asn	Cys	Ser	Val
				405					410					415	
Ile	Gln	Thr	Ser	Cys	Gln	Glu	Gly	Asp	Gln	Arg	Cys	Leu	Tyr	Asn	Pro
			420					425					430		
Cys	Leu	Asp	Ser	Cys	Gly	Gly	Ser	Ser	Leu	Cys	Asp	Pro	Asn	Asn	Ser
		435					440					445			
Leu	Asn	Asn	Cys	Ser	Gln	Cys	Glu	Pro	Ile	Thr	Leu	Glu	Leu	Cys	Met
	450					455						460			
Asn	Leu	Pro	Tyr	Asn	Ser	Thr	Ser	Tyr	Pro	Asn	Tyr	Phe	Gly	His	Arg
465					470					475					480
Thr	Gln	Lys	Glu	Ala	Ser	Ile	Ser	Trp	Glu	Ser	Ser	Leu	Phe	Pro	Ala
				485					490					495	
Leu	Val	Gln	Thr	Asn	Cys	Tyr	Lys	Tyr	Leu	Met	Phe	Phe	Ser	Cys	Thr
			500					505					510		
Ile	Leu	Val	Pro	Lys	Cys	Asp	Val	Asn	Thr	Gly	Glu	Arg	Ile	Pro	Pro
		515					520					525			
Cys	Arg	Ala	Leu	Cys	Glu	His	Ser	Lys	Glu	Arg	Cys	Glu	Ser	Val	Leu
	530					535					540				
Gly	Ile	Val	Gly	Leu	Gln	Trp	Pro	Glu	Asp	Thr	Asp	Cys	Ser	Gln	Phe

-59-

545					550					555				560
Pro	Glu	Glu	Asn	Ser	Asp	Asn	Gln	Thr	Cys	Leu	Met	Pro	Asp	Glu Tyr
				565					570					575
Val	Glu	Glu	Cys	Ser	Pro	Ser	His	Phe	Lys	Cys	Arg	Ser	Gly	Gln Cys
			580					585					590	
Val	Leu	Ala	Ser	Arg	Arg	Cys	Asp	Gly	Gln	Ala	Asp	Cys	Asp	Asp Asp
		595					600				605			
Ser	Asp	Glu	Glu	Asn	Cys	Gly	Cys	Lys	Glu	Arg	Asp	Leu	Trp	Glu Cys
	610					615					620			
Pro	Ser	Asn	Lys	Gln	Cys	Leu	Lys	His	Thr	Val	Ile	Cys	Asp	Gly Phe
625					630					635				640
Pro	Asp	Cys	Pro	Asp	Tyr	Met	Asp	Glu	Lys	Asn	Cys	Ser	Phe	Cys Gln
				645				650					655	
Asp	Asp	Glu	Leu	Glu	Cys	Ala	Asn	His	Ala	Cys	Val	Ser	Arg	Asp Leu
			660					665					670	
Trp	Cys	Asp	Gly	Glu	Ala	Asp	Cys	Ser	Asp	Ser	Ser	Asp	Glu	Trp Asp
		675					680					685		
Cys	Val	Thr	Leu	Ser	Ile	Asn	Val	Asn	Ser	Ser	Ser	Phe	Leu	Met Val
	690					695					700			
His	Arg	Ala	Ala	Thr	Glu	His	His	Val	Cys	Ala	Asp	Gly	Trp	Gln Glu
705					710					715				720
Ile	Leu	Ser	Gln	Leu	Ala	Cys	Lys	Gln	Met	Gly	Leu	Gly	Glu	Pro Ser
			725						730					735
Val	Thr	Lys	Leu	Ile	Gln	Glu	Gln	Glu	Lys	Glu	Pro	Arg	Trp	Leu Thr
			740					745					750	
Leu	His	Ser	Asn	Trp	Glu	Ser	Leu	Asn	Gly	Thr	Thr	Leu	His	Glu Leu
		755					760					765		
Leu	Val	Asn	Gly	Gln	Ser	Cys	Glu	Ser	Arg	Ser	Lys	Ile	Ser	Leu Leu
	770					775					780			
Cys	Thr	Lys	Gln	Asp	Cys	Gly	Arg	Arg	Pro	Ala	Ala	Arg	Met	Asn Lys
785				790					795					800
Arg	Ile	Leu	Gly	Gly	Arg	Thr	Ser	Arg	Pro	Gly	Arg	Trp	Pro	Trp Gln
			805						810					815
Cys	Ser	Leu	Gln	Ser	Glu	Pro	Ser	Gly	His	Ile	Cys	Gly	Cys	Val Leu
			820					825					830	
Ile	Ala	Lys	Lys	Trp	Val	Leu	Thr	Val	Ala	His	Cys	Phe	Glu	Gly Arg
		835					840					845		
Glu	Asn	Ala	Ala	Val	Trp	Lys	Val	Val	Leu	Gly	Ile	Asn	Asn	Leu Asp
	850					855					860			
His	Pro	Ser	Val	Phe	Met	Gln	Thr	Arg	Phe	Val	Lys	Thr	Ile	Ile Leu
865					870				875					880
His	Pro	Arg	Tyr	Ser	Arg	Ala	Val	Val	Asp	Tyr	Asp	Ile	Ser	Ile Val
				885					890					895
Glu	Leu	Ser	Glu	Asp	Ile	Ser	Glu	Thr	Gly	Tyr	Val	Arg	Pro	Val Cys
			900					905					910	
Leu	Pro	Asn	Pro	Glu	Gln	Trp	Leu	Glu	Pro	Asp	Thr	Tyr	Cys	Tyr Ile
		915					920					925		
Thr	Gly	Trp	Gly	His	Met	Gly	Asn	Lys	Met	Pro	Phe	Lys	Leu	Gln Glu
	930					935					940			
Gly	Glu	Val	Arg	Ile	Ile	Ser	Leu	Glu	His	Cys	Gln	Ser	Tyr	Phe Asp
945					950					955				960
Met	Lys	Thr	Ile	Thr	Thr	Arg	Met	Ile	Cys	Ala	Gly	Tyr	Glu	Ser Gly
				965					970					975
Thr	Val	Asp	Ser	Cys	Met	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys Glu
			980					985					990	
Lys	Pro	Gly	Gly	Arg	Trp	Thr	Leu	Phe	Gly	Leu	Thr	Ser	Trp	Gly Ser
		995					1000					1005		
Val	Cys	Phe	Ser	Lys	Val	Leu	Gly	Pro	Gly	Val	Tyr	Ser	Asn	Val Ser
	1010					1015					1020			
Tyr	Phe	Val	Glu	Trp	Ile	Lys	Arg	Gln	Ile	Tyr	Ile	Gln	Thr	Phe Leu
1025					1030				1035					1040

<300>  
<308> GenBank HSU09860  
<309> 1995-06-03

<400> 30																
accagacagt tcttaaatta gcaagccttc aaaaccaaaa atg ggg tcg aaa aga																55
Met Gly Ser Lys Arg																
1 5																
ggc ata tct tct agg cat cat tct ctc agc tcc tat gaa atc atg ttt																103
Gly Ile Ser Ser Arg His His Ser Leu Ser Ser Tyr Glu Ile Met Phe																
10 15 20																
gca gct ctc ttt gcc ata ttg gta gtg ctc tgt gct gga tta att gca																151
Ala Ala Leu Phe Ala Ile Leu Val Val Leu Cys Ala Gly Leu Ile Ala																
25 30 35																
gta tcc tgc ctg aca atc aag gaa tcc caa cga ggt gca gca ctt gga																199
Val Ser Cys Leu Thr Ile Lys Glu Ser Gln Arg Gly Ala Ala Leu Gly																
40 45 50																
cag agt cat gaa gcc aga gcg aca ttt aaa ata aca tcc gga gtt aca																247
Gln Ser His Glu Ala Arg Ala Thr Phe Lys Ile Thr Ser Gly Val Thr																
55 60 65																
tat aat cct aat ttg caa gac aaa ctc tca gtg gat ttc aaa gtt ctt																295
Tyr Asn Pro Asn Leu Gln Asp Lys Leu Ser Val Asp Phe Lys Val Leu																
70 75 80 85																
gct ttt gac ctt cag caa atg ata gat gag atc ttt cta tca agc aat																343
Ala Phe Asp Leu Gln Gln Met Ile Asp Glu Ile Phe Leu Ser Ser Asn																
90 95 100																
ctg aag aat gaa tat aag aac tca aga gtt tta caa ttt gaa aat ggc																391
Leu Lys Asn Glu Tyr Lys Asn Ser Arg Val Leu Gln Phe Glu Asn Gly																
105 110 115																
agc att ata gtc gta ttt gac ctt ttc ttt gcc cag tgg gtg tca gat																439
Ser Ile Ile Val Val Phe Asp Leu Phe Phe Ala Gln Trp Val Ser Asp																
120 125 130																
caa aat gta aaa gaa gaa ctg att caa ggc ctt gaa gca aat aaa tcc																487
Gln Asn Val Lys Glu Glu Leu Ile Gln Gly Leu Glu Ala Asn Lys Ser																
135 140 145																
agc caa ctg gtc act ttc cat att gat ttg aac agc gtt gat atc cta																535
Ser Gln Leu Val Thr Phe His Ile Asp Leu Asn Ser Val Asp Ile Leu																
150 155 160 165																

-61-

gac aag cta aca acc acc agt cat ctg gca act cca gga aat gtc tca Asp Lys Leu Thr Thr Thr Ser His Leu Ala Thr Pro Gly Asn Val Ser 170 175 180	583
ata gag tgc ctg cct ggt tca agt cct tgt act gat gct cta acg tgt Ile Glu Cys Leu Pro Gly Ser Ser Pro Cys Thr Asp Ala Leu Thr Cys 185 190 195	631
ata aaa gct gat tta ttt tgt gat gga gaa gta aac tgt cca gat ggt Ile Lys Ala Asp Leu Phe Cys Asp Gly Glu Val Asn Cys Pro Asp Gly 200 205 210	679
tct gac gaa gac aat aaa atg tgt gcc aca gtt tgt gat gga aga ttt Ser Asp Glu Asp Asn Lys Met Cys Ala Thr Val Cys Asp Gly Arg Phe 215 220 225	727
ttg tta act gga tca tct ggg tct ttc cag gct act cat tat cca aaa Leu Leu Thr Gly Ser Ser Gly Ser Phe Gln Ala Thr His Tyr Pro Lys 230 235 240 245	775
cct tct gaa aca agt gtt gtc tgc cag tgg atc ata cgt gta aac caa Pro Ser Glu Thr Ser Val Val Cys Gln Trp Ile Ile Arg Val Asn Gln 250 255 260	823
gga ctt tcc att aaa ctg agc ttc gat gat ttt aat aca tat tat aca Gly Leu Ser Ile Lys Leu Ser Phe Asp Asp Phe Asn Thr Tyr Tyr Thr 265 270 275	871
gat ata tta gat att tat gaa ggt gta gga tca agc aag att tta aga Asp Ile Leu Asp Ile Tyr Glu Gly Val Gly Ser Ser Lys Ile Leu Arg 280 285 290	919
gct tct att tgg gaa act aat cct ggc aca ata aga att ttt tcc aac Ala Ser Ile Trp Glu Thr Asn Pro Gly Thr Ile Arg Ile Phe Ser Asn 295 300 305	967
caa gtt act gcc acc ttt ctt ata gaa tct gat gaa agt gat tat gtt Gln Val Thr Ala Thr Phe Leu Ile Glu Ser Asp Glu Ser Asp Tyr Val 310 315 320 325	1015
ggc ttt aat gca aca tat act gca ttt aac agc agt gag ctt aat aat Gly Phe Asn Ala Thr Tyr Thr Ala Phe Asn Ser Ser Glu Leu Asn Asn 330 335 340	1063
tat gag aaa att aat tgt aac ttt gag gat ggc ttt tgt ttc tgg gtc Tyr Glu Lys Ile Asn Cys Asn Phe Glu Asp Gly Phe Cys Phe Trp Val 345 350 355	1111
cag gat cta aat gat gat aat gaa tgg gaa agg att cag gga agc acc Gln Asp Leu Asn Asp Asp Asn Glu Trp Glu Arg Ile Gln Gly Ser Thr 360 365 370	1159
ttt tct cct ttt act gga ccc aat ttt gac cac act ttt ggc aat gct Phe Ser Pro Phe Thr Gly Pro Asn Phe Asp His Thr Phe Gly Asn Ala 375 380 385	1207
tca gga ttt tac att tct acc cca act gga cca gga ggg aga caa gaa Ser Gly Phe Tyr Ile Ser Thr Pro Thr Gly Pro Gly Gly Arg Gln Glu 390 395 400 405	1255



-62-

cga	gtg	ggg	ctt	tta	agc	ctc	cct	ttg	gac	ccc	act	ttg	gag	cca	gct	1303
Arg	Val	Gly	Leu	Leu	Ser	Leu	Pro	Leu	Asp	Pro	Thr	Leu	Glu	Pro	Ala	
			410						415					420		
tgc	ctt	agt	ttc	tgg	tat	cat	atg	tat	ggt	gaa	aat	gtc	cat	aaa	tta	1351
Cys	Leu	Ser	Phe	Trp	Tyr	His	Met	Tyr	Gly	Glu	Asn	Val	His	Lys	Leu	
			425					430					435			
agc	att	aat	atc	agc	aat	gac	caa	aat	atg	gag	aag	aca	gtt	ttc	caa	1399
Ser	Ile	Asn	Ile	Ser	Asn	Asp	Gln	Asn	Met	Glu	Lys	Thr	Val	Phe	Gln	
		440					445					450				
aag	gaa	gga	aat	tat	gga	gac	aat	tgg	aat	tat	gga	caa	gta	acc	cta	1447
Lys	Glu	Gly	Asn	Tyr	Gly	Asp	Asn	Trp	Asn	Tyr	Gly	Gln	Val	Thr	Leu	
	455					460					465					
aat	gaa	aca	gtt	aaa	ttt	aag	gtt	gct	ttt	aat	gct	ttt	aaa	aac	aag	1495
Asn	Glu	Thr	Val	Lys	Phe	Lys	Val	Ala	Phe	Asn	Ala	Phe	Lys	Asn	Lys	
470					475					480					485	
atc	ctg	agt	gat	att	gcg	ttg	gat	gac	att	agc	cta	aca	tat	ggg	att	1543
Ile	Leu	Ser	Asp	Ile	Ala	Leu	Asp	Asp	Ile	Ser	Leu	Thr	Tyr	Gly	Ile	
				490					495					500		
tgc	aat	ggg	agt	ctt	tat	cca	gaa	cca	act	ttg	gtg	cca	act	cct	cca	1591
Cys	Asn	Gly	Ser	Leu	Tyr	Pro	Glu	Pro	Thr	Leu	Val	Pro	Thr	Pro	Pro	
			505					510					515			
cca	gaa	ctt	cct	acg	gac	tgt	gga	gga	cct	ttt	gag	ctg	tgg	gag	cca	1639
Pro	Glu	Leu	Pro	Thr	Asp	Cys	Gly	Gly	Pro	Phe	Glu	Leu	Trp	Glu	Pro	
		520					525					530				
aat	aca	aca	ttc	agt	tct	acg	aac	ttt	cca	aac	agc	tac	cct	aat	ctg	1687
Asn	Thr	Thr	Phe	Ser	Ser	Thr	Asn	Phe	Pro	Asn	Ser	Tyr	Pro	Asn	Leu	
	535					540					545					
gct	ttc	tgt	gtt	tgg	att	tta	aat	gca	caa	aaa	gga	aag	aat	ata	caa	1735
Ala	Phe	Cys	Val	Trp	Ile	Leu	Asn	Ala	Gln	Lys	Gly	Lys	Asn	Ile	Gln	
550					555					560					565	
ctt	cat	ttt	caa	gaa	ttt	gac	tta	gaa	aat	att	aac	gat	gta	gtt	gaa	1783
Leu	His	Phe	Gln	Glu	Phe	Asp	Leu	Glu	Asn	Ile	Asn	Asp	Val	Val	Glu	
				570					575					580		
ata	aga	gat	ggt	gaa	gaa	gct	gat	tcc	ttg	ctc	tta	gct	gtg	tac	aca	1831
Ile	Arg	Asp	Gly	Glu	Glu	Ala	Asp	Ser	Leu	Leu	Leu	Ala	Val	Tyr	Thr	
			585					590					595			
ggg	cct	ggc	cca	gta	aag	gat	gtg	ttc	tct	acc	acc	aac	aga	atg	act	1879
Gly	Pro	Gly	Pro	Val	Lys	Asp	Val	Phe	Ser	Thr	Thr	Asn	Arg	Met	Thr	
		600					605					610				
gtg	ctt	ctc	atc	act	aac	gat	gtg	ttg	gca	aga	gga	ggg	ttt	aaa	gca	1927
Val	Leu	Leu	Ile	Thr	Asn	Asp	Val	Leu	Ala	Arg	Gly	Gly	Phe	Lys	Ala	
	615					620					625					
aac	ttt	act	act	ggc	tat	cac	ttg	ggg	att	cca	gag	cca	tgc	aag	gca	1975
Asn	Phe	Thr	Thr	Gly	Tyr	His	Leu	Gly	Ile	Pro	Glu	Pro	Cys	Lys	Ala	
630					635					640					645	
gac	cat	ttt	caa	tgt	aaa	aat	gga	gag	tgt	gtt	cca	ctg	gtg	aat	ctc	2023

-63-

Asp	His	Phe	Gln	Cys	Lys	Asn	Gly	Glu	Cys	Val	Pro	Leu	Val	Asn	Leu	
				650					655					660		
tgt	gac	ggc	cat	ctg	cac	tgt	gag	gat	ggc	tca	gat	gaa	gca	gat	tgt	2071
Cys	Asp	Gly	His	Leu	His	Cys	Glu	Asp	Gly	Ser	Asp	Glu	Ala	Asp	Cys	
			665				670					675				
gtg	cgt	ttt	ttc	aat	ggc	aca	acg	aac	aac	aat	ggc	tta	gtg	cgg	ttc	2119
Val	Arg	Phe	Phe	Asn	Gly	Thr	Thr	Asn	Asn	Asn	Gly	Leu	Val	Arg	Phe	
		680					685					690				
aga	atc	cag	agc	ata	tgg	cat	aca	gct	tgt	gct	gag	aac	tgg	acc	acc	2167
Arg	Ile	Gln	Ser	Ile	Trp	His	Thr	Ala	Cys	Ala	Glu	Asn	Trp	Thr	Thr	
	695					700					705					
cag	att	tca	aat	gat	ggt	tgt	caa	ctg	ctg	gga	cta	ggg	agt	gga	aac	2215
Gln	Ile	Ser	Asn	Asp	Val	Cys	Gln	Leu	Leu	Gly	Leu	Gly	Ser	Gly	Asn	
710					715				720						725	
tca	tca	aag	cca	atc	ttc	tct	acc	gat	ggc	gga	cca	ttt	gtc	aaa	tta	2263
Ser	Ser	Lys	Pro	Ile	Phe	Ser	Thr	Asp	Gly	Gly	Pro	Phe	Val	Lys	Leu	
				730					735				740			
aac	aca	gca	cct	gat	ggc	cac	tta	ata	cta	aca	ccc	agt	caa	cag	tgt	2311
Asn	Thr	Ala	Pro	Asp	Gly	His	Leu	Ile	Leu	Thr	Pro	Ser	Gln	Gln	Cys	
			745				750						755			
tta	cag	gat	tcc	ttg	att	cgg	tta	cag	tgt	aac	cat	aaa	tct	tgt	gga	2359
Leu	Gln	Asp	Ser	Leu	Ile	Arg	Leu	Gln	Cys	Asn	His	Lys	Ser	Cys	Gly	
	760					765						770				
aaa	aaa	ctg	gca	gct	caa	gac	atc	acc	cca	aag	att	gtt	gga	gga	agt	2407
Lys	Lys	Leu	Ala	Ala	Gln	Asp	Ile	Thr	Pro	Lys	Ile	Val	Gly	Gly	Ser	
	775					780					785					
aat	gcc	aaa	gaa	ggg	gcc	tgg	ccc	tgg	gtt	gtg	ggc	ctg	tat	tat	ggc	2455
Asn	Ala	Lys	Glu	Gly	Ala	Trp	Pro	Trp	Val	Val	Gly	Leu	Tyr	Tyr	Gly	
790					795				800						805	
ggc	cga	ctg	ctc	tgc	ggc	gca	tct	ctc	gtc	agc	agt	gac	tgg	ctg	gtg	2503
Gly	Arg	Leu	Leu	Cys	Gly	Ala	Ser	Leu	Val	Ser	Ser	Asp	Trp	Leu	Val	
				810					815				820			
tcc	gcc	gca	cac	tgc	gtg	tat	ggg	aga	aac	tta	gag	cca	tcc	aag	tgg	2551
Ser	Ala	Ala	His	Cys	Val	Tyr	Gly	Arg	Asn	Leu	Glu	Pro	Ser	Lys	Trp	
			825				830						835			
aca	gca	atc	cta	ggc	ctg	cat	atg	aaa	tca	aat	ctg	acc	tct	cct	caa	2599
Thr	Ala	Ile	Leu	Gly	Leu	His	Met	Lys	Ser	Asn	Leu	Thr	Ser	Pro	Gln	
		840					845					850				
aca	gtc	cct	cga	tta	ata	gat	gaa	att	gtc	ata	aac	cct	cat	tac	aat	2647
Thr	Val	Pro	Arg	Leu	Ile	Asp	Glu	Ile	Val	Ile	Asn	Pro	His	Tyr	Asn	
	855					860					865					
agg	cga	aga	aag	gac	aac	gac	att	gcc	atg	atg	cat	ctg	gaa	ttt	aaa	2695
Arg	Arg	Arg	Lys	Asp	Asn	Asp	Ile	Ala	Met	Met	His	Leu	Glu	Phe	Lys	
870					875					880					885	
gtg	aat	tac	aca	gat	tac	ata	caa	cct	att	tgt	tta	ccg	gaa	gaa	aat	2743
Val	Asn	Tyr	Thr	Asp	Tyr	Ile	Gln	Pro	Ile	Cys	Leu	Pro	Glu	Glu	Asn	

-64-

890										895					900					
caa gtt ttt cct cca gga aga aat tgt tct att gct ggt tgg ggg acg	2791																			
Gln Val Phe Pro Pro Gly Arg Asn Cys Ser Ile Ala Gly Trp Gly Thr																				
905 910 915																				
gtt gta tat caa ggt act act gca aac ata ttg caa gaa gct gat gtt	2839																			
Val Val Tyr Gln Gly Thr Thr Ala Asn Ile Leu Gln Glu Ala Asp Val																				
920 925 930																				
cct ctt cta tca aat gag aga tgc caa cag cag atg cca gaa tat aac	2887																			
Pro Leu Leu Ser Asn Glu Arg Cys Gln Gln Gln Met Pro Glu Tyr Asn																				
935 940 945																				
att act gaa aat atg ata tgt gca ggc tat gaa gaa gga gga ata gat	2935																			
Ile Thr Glu Asn Met Ile Cys Ala Gly Tyr Glu Glu Gly Gly Ile Asp																				
950 955 960 965																				
tct tgt cag ggg gat tca gga gga cca tta atg tgc caa gaa aac aac	2983																			
Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu Asn Asn																				
970 975 980																				
agg tgg ttc ctt gct ggt gtg acc tca ttt gga tac aag tgt gcc ctg	3031																			
Arg Trp Phe Leu Ala Gly Val Thr Ser Phe Gly Tyr Lys Cys Ala Leu																				
985 990 995																				
cct aat cgc ccc gga gtg tat gcc agg gtc tca agg ttt acc gaa tgg	3079																			
Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Ser Arg Phe Thr Glu Trp																				
1000 1005 1010																				
ata caa agt ttt cta cat tag cgcatttctt aaactaaaca ggaaagtctgc	3130																			
Ile Gln Ser Phe Leu His *																				
1015																				
attattttcc cattctactc tagaaagcat ggaaattaag tgtttcgtac aaaaatttta	3190																			
aaaagttacc aaagggtttt attcttacct atgtcaatga aatgctaggg ggccagggaa	3250																			
acaaaatttt aaaaataata aaattcacca tagcaataga gaataacttt aaaataccat	3310																			
taaatacatt tgtatttcat tgtgaacagg tatttcttca cagatctcat ttttaaaatt	3370																			
cttaatgatt atttttatta cttactgttg ttttaaggga tgttatttta aagcatatac	3430																			
catacactta agaaatttga gcagaattta aaaaagaaag aaaataaatt gtttttccca	3490																			
aagtatgtca ctggttgaaa taaactgcca taaattttct agttccagtt tagtttgctg	3550																			
ctattagcag aaactcaatt gtttctctgt cttttctatc aaaattttca acatatgcat	3610																			
aaccttagta ttttcccaac caatagaaac tattttattgt aagcttatgt cacaggcctg	3670																			
gactaaattg attttacgtt cctctt	3696																			
<210> 31																				
<211> 1019																				
<212> PRT																				
<213> Homo Sapien																				
<400> 31																				
Met Gly Ser Lys Arg Gly Ile Ser Ser Arg His His Ser Leu Ser Ser																				
1 5 10 15																				
Tyr Glu Ile Met Phe Ala Ala Leu Phe Ala Ile Leu Val Val Leu Cys																				
20 25 30																				
Ala Gly Leu Ile Ala Val Ser Cys Leu Thr Ile Lys Glu Ser Gln Arg																				
35 40 45																				
Gly Ala Ala Leu Gly Gln Ser His Glu Ala Arg Ala Thr Phe Lys Ile																				
50 55 60																				
Thr Ser Gly Val Thr Tyr Asn Pro Asn Leu Gln Asp Lys Leu Ser Val																				
65 70 75 80																				

-65-

Asp	Phe	Lys	Val	Leu	Ala	Phe	Asp	Leu	Gln	Gln	Met	Ile	Asp	Glu	Ile
				85					90					95	
Phe	Leu	Ser	Ser	Asn	Leu	Lys	Asn	Glu	Tyr	Lys	Asn	Ser	Arg	Val	Leu
			100					105					110		
Gln	Phe	Glu	Asn	Gly	Ser	Ile	Ile	Val	Val	Phe	Asp	Leu	Phe	Phe	Ala
		115					120					125			
Gln	Trp	Val	Ser	Asp	Gln	Asn	Val	Lys	Glu	Glu	Leu	Ile	Gln	Gly	Leu
	130					135					140				
Glu	Ala	Asn	Lys	Ser	Ser	Gln	Leu	Val	Thr	Phe	His	Ile	Asp	Leu	Asn
145					150					155					160
Ser	Val	Asp	Ile	Leu	Asp	Lys	Leu	Thr	Thr	Thr	Ser	His	Leu	Ala	Thr
				165					170					175	
Pro	Gly	Asn	Val	Ser	Ile	Glu	Cys	Leu	Pro	Gly	Ser	Ser	Pro	Cys	Thr
			180					185					190		
Asp	Ala	Leu	Thr	Cys	Ile	Lys	Ala	Asp	Leu	Phe	Cys	Asp	Gly	Glu	Val
	195						200					205			
Asn	Cys	Pro	Asp	Gly	Ser	Asp	Glu	Asp	Asn	Lys	Met	Cys	Ala	Thr	Val
	210					215					220				
Cys	Asp	Gly	Arg	Phe	Leu	Leu	Thr	Gly	Ser	Ser	Gly	Ser	Phe	Gln	Ala
225					230					235					240
Thr	His	Tyr	Pro	Lys	Pro	Ser	Glu	Thr	Ser	Val	Val	Cys	Gln	Trp	Ile
				245					250					255	
Ile	Arg	Val	Asn	Gln	Gly	Leu	Ser	Ile	Lys	Leu	Ser	Phe	Asp	Asp	Phe
			260					265					270		
Asn	Thr	Tyr	Tyr	Thr	Asp	Ile	Leu	Asp	Ile	Tyr	Glu	Gly	Val	Gly	Ser
		275					280					285			
Ser	Lys	Ile	Leu	Arg	Ala	Ser	Ile	Trp	Glu	Thr	Asn	Pro	Gly	Thr	Ile
	290					295					300				
Arg	Ile	Phe	Ser	Asn	Gln	Val	Thr	Ala	Thr	Phe	Leu	Ile	Glu	Ser	Asp
305					310					315					320
Glu	Ser	Asp	Tyr	Val	Gly	Phe	Asn	Ala	Thr	Tyr	Thr	Ala	Phe	Asn	Ser
				325					330					335	
Ser	Glu	Leu	Asn	Asn	Tyr	Glu	Lys	Ile	Asn	Cys	Asn	Phe	Glu	Asp	Gly
			340					345					350		
Phe	Cys	Phe	Trp	Val	Gln	Asp	Leu	Asn	Asp	Asp	Asn	Glu	Trp	Glu	Arg
		355					360					365			
Ile	Gln	Gly	Ser	Thr	Phe	Ser	Pro	Phe	Thr	Gly	Pro	Asn	Phe	Asp	His
	370					375					380				
Thr	Phe	Gly	Asn	Ala	Ser	Gly	Phe	Tyr	Ile	Ser	Thr	Pro	Thr	Gly	Pro
385					390					395					400
Gly	Gly	Arg	Gln	Glu	Arg	Val	Gly	Leu	Leu	Ser	Leu	Pro	Leu	Asp	Pro
				405					410					415	
Thr	Leu	Glu	Pro	Ala	Cys	Leu	Ser	Phe	Trp	Tyr	His	Met	Tyr	Gly	Glu
			420					425					430		
Asn	Val	His	Lys	Leu	Ser	Ile	Asn	Ile	Ser	Asn	Asp	Gln	Asn	Met	Glu
		435					440					445			
Lys	Thr	Val	Phe	Gln	Lys	Glu	Gly	Asn	Tyr	Gly	Asp	Asn	Trp	Asn	Tyr
	450					455					460				
Gly	Gln	Val	Thr	Leu	Asn	Glu	Thr	Val	Lys	Phe	Lys	Val	Ala	Phe	Asn
465					470					475					480
Ala	Phe	Lys	Asn	Lys	Ile	Leu	Ser	Asp	Ile	Ala	Leu	Asp	Asp	Ile	Ser
				485					490					495	
Leu	Thr	Tyr	Gly	Ile	Cys	Asn	Gly	Ser	Leu	Tyr	Pro	Glu	Pro	Thr	Leu
			500					505					510		
Val	Pro	Thr	Pro	Pro	Pro	Glu	Leu	Pro	Thr	Asp	Cys	Gly	Gly	Pro	Phe
		515					520					525			
Glu	Leu	Trp	Glu	Pro	Asn	Thr	Thr	Phe	Ser	Ser	Thr	Asn	Phe	Pro	Asn
	530					535					540				
Ser	Tyr	Pro	Asn	Leu	Ala	Phe	Cys	Val	Trp	Ile	Leu	Asn	Ala	Gln	Lys
545					550					555					560
Gly	Lys	Asn	Ile	Gln	Leu	His	Phe	Gln	Glu	Phe	Asp	Leu	Glu	Asn	Ile

-66-

Asn	Asp	Val	Val	565	Glu	Ile	Arg	Asp	Gly	570	Glu	Glu	Ala	Asp	Ser	575	Leu	Leu
Leu	Ala	Val	Tyr	580	Thr	Gly	Pro	Gly	585	Pro	Val	Lys	Asp	Val	Phe	590	Ser	Thr
Thr	Asn	Arg	Met	595	Thr	Val	Leu	Leu	600	Ile	Thr	Asn	Asp	Val	Leu	605	Ala	Arg
Gly	Gly	Phe	Lys	610	Ala	Asn	Phe	Thr	615	Thr	Gly	Tyr	His	Leu	Gly	620	Ile	Pro
625	Glu	Pro	Cys	630	Lys	Ala	Asp	His	635	Phe	Gln	Cys	Lys	Asn	Gly	640	Glu	Cys
645	Pro	Leu	Val	650	Asn	Leu	Cys	Asp	655	Gly	His	Leu	His	Cys	Glu	660	Asp	Gly
665	Asp	Glu	Ala	670	Asp	Cys	Val	Arg	675	Phe	Phe	Asn	Gly	Thr	Thr	680	Asn	Asn
685	Gly	Leu	Val	690	Arg	Phe	Arg	Ile	695	Gln	Ser	Ile	Trp	His	Thr	700	Ala	Cys
705	Glu	Asn	Trp	710	Thr	Thr	Gln	Ile	715	Ser	Asn	Asp	Val	Cys	Gln	720	Leu	Leu
725	Leu	Gly	Ser	730	Gly	Asn	Ser	Ser	735	Lys	Pro	Ile	Phe	Ser	Thr	740	Asp	Gly
745	Pro	Phe	Val	750	Lys	Leu	Asn	Thr	755	Ala	Pro	Asp	Gly	His	Leu	760	Ile	Leu
765	Pro	Ser	Gln	770	Gln	Cys	Leu	Gln	775	Asp	Ser	Leu	Ile	Arg	Leu	780	Gln	Cys
785	His	Lys	Ser	790	Cys	Gly	Lys	Lys	795	Leu	Ala	Ala	Gln	Asp	Ile	795	Thr	Pro
800	Ile	Val	Gly	805	Gly	Ser	Asn	Ala	810	Lys	Glu	Gly	Ala	Trp	Pro	815	Trp	Val
820	Gly	Leu	Tyr	825	Tyr	Gly	Gly	Arg	830	Leu	Leu	Cys	Gly	Ala	Ser	835	Leu	Val
840	Ser	Asp	Trp	845	Leu	Val	Ser	Ala	850	Ala	His	Cys	Val	Tyr	Gly	855	Arg	Asn
860	Glu	Pro	Ser	865	Lys	Trp	Thr	Ala	870	Ile	Leu	Gly	Leu	His	Met	875	Lys	Ser
880	Leu	Thr	Ser	885	Pro	Gln	Thr	Val	890	Pro	Arg	Leu	Ile	Asp	Glu	895	Ile	Val
900	Asn	Pro	His	905	Tyr	Asn	Arg	Arg	910	Lys	Asp	Asn	Asp	Ile	Ala	915	Met	Met
920	865	Leu	Glu	925	Phe	Lys	Val	Asn	930	Thr	Asp	Tyr	Ile	Gln	Pro	935	Ile	Cys
940	His	Pro	Glu	945	Glu	Asn	Gln	Val	950	Phe	Pro	Pro	Gly	Arg	Asn	955	Cys	Ser
960	Leu	Pro	Glu	965	Thr	Val	Val	Tyr	970	Gln	Gly	Thr	Thr	Ala	Asn	975	Ile	Leu
975	Gln	Glu	Ala	980	Asp	Val	Pro	Leu	985	Leu	Ser	Asn	Glu	Arg	Cys	990	Gln	Gln
990	Met	Pro	Glu	995	Tyr	Asn	Ile	Thr	1000	Glu	Asn	Met	Ile	Cys	Ala	1005	Gly	Tyr
1010	Glu	Gly	Gly	1015	Ile	Asp	Ser	Cys	1020	Gln	Gly	Asp	Ser	Gly	Gly	1025	Pro	Leu
1030	Cys	Gln	Glu	1035	Asn	Asn	Arg	Trp	1040	Phe	Leu	Ala	Gly	Val	Thr	1045	Ser	Phe
1050	Tyr	Lys	Cys	1055	Ala	Leu	Pro	Asn	1060	Arg	Pro	Gly	Val	Tyr	Ala	1065	Arg	Val
1070	Arg	Phe	Thr	1075	Glu	Trp	Ile	Gln	1080	Ser	Phe	Leu	His			1085		Ser

<210> 32  
 <211> 1500  
 <212> DNA

-67-

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (62)...(1318)

<223> Nucleotide sequence encoding human airway  
trypsin-like protease

&lt;300&gt;

&lt;308&gt; GenBank AB002134

&lt;309&gt; 1998-06-04

&lt;400&gt; 32

```

gagtgggaat ctcaaagcag ttgagtaggc agaaaaaaga acctcttcat taaggattaa      60
a atg tat agg cca gca cgt gta act tcg act tca aga ttt ctg aat cca      109
  Met Tyr Arg Pro Ala Arg Val Thr Ser Thr Ser Arg Phe Leu Asn Pro
    1             5             10             15

tat gta gta tgt ttc att gtc gtc gca ggg gta gtg atc ctg gca gtc      157
Tyr Val Val Cys Phe Ile Val Val Ala Gly Val Val Ile Leu Ala Val
          20             25             30

acc ata gct cta ctt gtt tac ttt tta gct ttt gat caa aaa tct tac      205
Thr Ile Ala Leu Leu Val Tyr Phe Leu Ala Phe Asp Gln Lys Ser Tyr
          35             40             45

ttt tat agg agc agt ttt caa ctc cta aat gtt gaa tat aat agt cag      253
Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln
          50             55             60

tta aat tca cca gct aca cag gaa tac agg act ttg agt gga aga att      301
Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile
          65             70             75             80

gaa tct ctg att act aaa aca ttc aaa gaa tca aat tta aga aat cag      349
Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln
          85             90             95

ttc atc aga gct cat gtt gcc aaa ctg agg caa gat ggt agt ggt gtg      397
Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val
          100            105            110

aga gcg gat gtt gtc atg aaa ttt caa ttc act aga aat aac aat gga      445
Arg Ala Asp Val Val Met Lys Phe Gln Phe Thr Arg Asn Asn Asn Gly
          115            120            125

gca tca atg aaa agc aga att gag tct gtt tta cga caa atg ctg aat      493
Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn
          130            135            140

aac tct gga aac ctg gaa ata aac cct tca act gag ata aca tca ctt      541
Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu
          145            150            155            160

act gac cag gct gca gca aat tgg ctt att aat gaa tgt ggg gcc ggt      589
Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly
          165            170            175

cca gac cta ata aca ttg tct gag cag aga atc ctt gga ggc act gag      637
Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu
          180            185            190

```

-68-

gct gag gag gga agc tgg ccg tgg caa gtc agt ctg cgg ctc aat aat	685
Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn	
195 200 205	
gcc cac cac tgt gga ggc agc ctg atc aat aac atg tgg atc ctg aca	733
Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr	
210 215 220	
gca gct cac tgc ttc aga agc aac tct aat cct cgt gac tgg att gcc	781
Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala	
225 230 235 240	
acg tct ggt att tcc aca aca ttt cct aaa cta aga atg aga gta aga	829
Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg	
245 250 255	
aat att tta att cat aac aat tat aaa tct gca act cat gaa aat gac	877
Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp	
260 265 270	
att gca ctt gtg aga ctt gag aac agt gtc acc ttt acc aaa gat atc	925
Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile	
275 280 285	
cat agt gtg tgt ctc cca gct gct acc cag aat att cca cct ggc tct	973
His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser	
290 295 300	
act gct tat gta aca gga tgg ggc gct caa gaa tat gct ggc cac aca	1021
Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr	
305 310 315 320	
gtt cca gag cta agg caa gga cag gtc aga ata ata agt aat gat gta	1069
Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val	
325 330 335	
tgt aat gca cca cat agt tat aat gga gcc atc ttg tct gga atg ctg	1117
Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu	
340 345 350	
tgt gct gga gta cct caa ggt gga gtg gac gca tgt cag ggt gac tct	1165
Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser	
355 360 365	
ggt ggc cca cta gta caa gaa gac tca cgg cgg ctt tgg ttt att gtg	1213
Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val	
370 375 380	
ggg ata gta agc tgg gga gat cag tgt ggc ctg ccg gat aag cca gga	1261
Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly	
385 390 395 400	
gtg tat act cga gtg aca gcc tac ctt gac tgg att agg caa caa act	1309
Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr	
405 410 415	
ggg atc tag tgcaacaagt gcacccctgt tgcaaagtct gtatgcaggt	1358
Gly Ile *	

-69-

```

gtgcctgtct taaattccaa agctttacat ttcaactgaa aaagaaacta gaaatgtcct 1418
aatttaacat cttgttacat aaatatggtt taacaaacac tgtttaacct ttctttatta 1478
ttaaagggtt tctattttct cc 1500

```

<210> 33  
 <211> 418  
 <212> PRT  
 <213> Homo Sapien

<400> 33  
 Met Tyr Arg Pro Ala Arg Val Thr Ser Thr Ser Arg Phe Leu Asn Pro  
 1 5 10 15  
 Tyr Val Val Cys Phe Ile Val Val Ala Gly Val Val Ile Leu Ala Val  
 20 25 30  
 Thr Ile Ala Leu Leu Val Tyr Phe Leu Ala Phe Asp Gln Lys Ser Tyr  
 35 40 45  
 Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln  
 50 55 60  
 Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile  
 65 70 75 80  
 Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln  
 85 90 95  
 Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val  
 100 105 110  
 Arg Ala Asp Val Val Met Lys Phe Gln Phe Thr Arg Asn Asn Asn Gly  
 115 120 125  
 Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn  
 130 135 140  
 Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu  
 145 150 155 160  
 Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly  
 165 170 175  
 Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu  
 180 185 190  
 Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn  
 195 200 205  
 Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr  
 210 215 220  
 Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala  
 225 230 235 240  
 Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg  
 245 250 255  
 Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp  
 260 265 270  
 Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile  
 275 280 285  
 His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser  
 290 295 300  
 Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr  
 305 310 315 320  
 Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val  
 325 330 335  
 Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu  
 340 345 350  
 Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser  
 355 360 365  
 Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val  
 370 375 380  
 Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly  
 385 390 395 400  
 Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr



-70-

Gly Ile 405 410 415  
 <210> 34  
 <211> 1783  
 <212> DNA  
 <213> Homo Sapien  
 <220>  
 <221> CDS  
 <222> (246)...(1499)  
 <223> Nucleic acid encoding human hepsin  
 <300>  
 <308> GenBank M18930  
 <309> 1993-06-11  
 <400> 34  
 tcgagcccgcc tttccaggga ccctacctga gggccacacag gtgaggcagc ctggcctagc 60  
 aggccccacg ccaccgcctc tgcctccagg ccgcccgcgtg ctgcggggcc accatgctcc 120  
 tgcccaggcc tggagactga cccgaccccg gcactacctc gaggtccgc cccacactgc 180  
 tggacccag ggtcccaccc tggcccagga ggtagccag ggaatcatta acaagaggca 240  
 gtgac atg gcg cag aag gag ggt ggc cgg act gtg cca tgc tgc tcc aga 290  
 Met Ala Gln Lys Glu Gly Gly Arg Thr Val Pro Cys Cys Ser Arg  
 1 5 10 15  
 ccc aag gtg gca gct ctc act gcg ggg acc ctg cta ctt ctg aca gcc 338  
 Pro Lys Val Ala Ala Leu Thr Ala Gly Thr Leu Leu Leu Leu Thr Ala  
 20 25 30  
 atc ggg gcg gca tcc tgg gcc att gtg gct gtt ctc ctc agg agt gac 386  
 Ile Gly Ala Ala Ser Trp Ala Ile Val Ala Val Leu Leu Arg Ser Asp  
 35 40 45  
 cag gag ccg ctg tac cca gtg cag gtc agc tct gcg gac gct cgg ctc 434  
 Gln Glu Pro Leu Tyr Pro Val Gln Val Ser Ser Ala Asp Ala Arg Leu  
 50 55 60  
 atg gtc ttt gac aag acg gaa ggg acg tgg cgg ctg ctg tgc tcc tcg 482  
 Met Val Phe Asp Lys Thr Glu Gly Thr Trp Arg Leu Leu Cys Ser Ser  
 65 70 75  
 cgc tcc aac gcc agg gta gcc gga ctc agc tgc gag gag atg ggc ttc 530  
 Arg Ser Asn Ala Arg Val Ala Gly Leu Ser Cys Glu Glu Met Gly Phe  
 80 85 90 95  
 ctc agg gca ctg acc cac tcc gag ctg gac gtg cga acg gcg ggc gcc 578  
 Leu Arg Ala Leu Thr His Ser Glu Leu Asp Val Arg Thr Ala Gly Ala  
 100 105 110  
 aat ggc acg tcg ggc ttc ttc tgt gtg gac gag ggg agg ctg ccc cac 626  
 Asn Gly Thr Ser Gly Phe Phe Cys Val Asp Glu Gly Arg Leu Pro His  
 115 120 125  
 acc cag agg ctg ctg gag gtc atc tcc gtg tgt gat tgc ccc aga ggc 674  
 Thr Gln Arg Leu Leu Glu Val Ile Ser Val Cys Asp Cys Pro Arg Gly  
 130 135 140  
 cgt ttc ttg gcc gcc atc tgc caa gac tgt ggc cgc agg aag ctg ccc 722  
 Arg Phe Leu Ala Ala Ile Cys Gln Asp Cys Gly Arg Arg Lys Leu Pro  
 145 150 155

-71-

gtg gac cgc atc gtg gga ggc cgg gac acc agc ttg ggc cgg tgg ccg Val Asp Arg Ile Val Gly Gly Arg Asp Thr Ser Leu Gly Arg Trp Pro 160 165 170 175	770
tgg caa gtc agc ctt cgc tat gat gga gca cac ctc tgt ggg gga tcc Trp Gln Val Ser Leu Arg Tyr Asp Gly Ala His Leu Cys Gly Gly Ser 180 185 190	818
ctg ctc tcc ggg gac tgg gtg ctg aca gcc gcc cac tgc ttc ccg gag Leu Leu Ser Gly Asp Trp Val Leu Thr Ala Ala His Cys Phe Pro Glu 195 200 205	866
cgg aac cgg gtc ctg tcc cga tgg cga gtg ttt gcc ggt gcc gtg gcc Arg Asn Arg Val Leu Ser Arg Trp Arg Val Phe Ala Gly Ala Val Ala 210 215 220	914
cag gcc tct ccc cac ggt ctg cag ctg ggg gtg cag gct gtg gtc tac Gln Ala Ser Pro His Gly Leu Gln Leu Gly Val Gln Ala Val Val Tyr 225 230 235	962
cac ggg ggc tat ctt ccc ttt cgg gac ccc aac agc gag gag aac agc His Gly Gly Tyr Leu Pro Phe Arg Asp Pro Asn Ser Glu Glu Asn Ser 240 245 250 255	1010
aac gat att gcc ctg gtc cac ctc tcc agt ccc ctg ccc ctc aca gaa Asn Asp Ile Ala Leu Val His Leu Ser Ser Pro Leu Pro Leu Thr Glu 260 265 270	1058
tac atc cag cct gtg tgc ctc cca gct gcc ggc cag gcc ctg gtg gat Tyr Ile Gln Pro Val Cys Leu Pro Ala Ala Gly Gln Ala Leu Val Asp 275 280 285	1106
ggc aag atc tgt acc gtg acg ggc tgg ggc aac acg cag tac tat ggc Gly Lys Ile Cys Thr Val Thr Gly Trp Gly Asn Thr Gln Tyr Tyr Gly 290 295 300	1154
caa cag gcc ggg gta ctc cag gag gct cga gtc ccc ata atc agc aat Gln Gln Ala Gly Val Leu Gln Glu Ala Arg Val Pro Ile Ile Ser Asn 305 310 315	1202
gat gtc tgc aat ggc gct gac ttc tat gga aac cag atc aag ccc aag Asp Val Cys Asn Gly Ala Asp Phe Tyr Gly Asn Gln Ile Lys Pro Lys 320 325 330 335	1250
atg ttc tgt gct ggc tac ccc gag ggt ggc att gat gcc tgc cag ggc Met Phe Cys Ala Gly Tyr Pro Glu Gly Gly Ile Asp Ala Cys Gln Gly 340 345 350	1298
gac agc ggt ggt ccc ttt gtg tgt gag gac agc atc tct cgg acg cca Asp Ser Gly Gly Pro Phe Val Cys Glu Asp Ser Ile Ser Arg Thr Pro 355 360 365	1346
cgt tgg cgg ctg tgt ggc att gtg agt tgg ggc act ggc tgt gcc ctg Arg Trp Arg Leu Cys Gly Ile Val Ser Trp Gly Thr Gly Cys Ala Leu 370 375 380	1394
gcc cag aag cca ggc gtc tac acc aaa gtc agt gac ttc cgg gag tgg Ala Gln Lys Pro Gly Val Tyr Thr Lys Val Ser Asp Phe Arg Glu Trp 385 390 395	1442

-72-

atc ttc cag gcc ata aag act cac tcc gaa gcc agc ggc atg gtg acc 1490  
 Ile Phe Gln Ala Ile Lys Thr His Ser Glu Ala Ser Gly Met Val Thr  
 400 405 410 415

cag ctc tga ccggtggctt ctgctgcgc agcctccagg gcccagggtg 1539  
 Gln Leu \*

atccccggtgg tgggatccac gctggggccga ggatgggacg tttttcttct tgggcccggg 1599  
 ccacaggtcc aaggacaccc tccctccagg gtctctctt ccacagtggc gggcccactc 1659  
 agccccgaga ccaccaacc tcacctcct gaccccatg taaatattgt tctgctgtct 1719  
 gggactcctg tctaggtgcc cctgatgatg ggatgctctt taaataataa agatgggttt 1779  
 gatt 1783

<210> 35  
 <211> 417  
 <212> PRT  
 <213> Homo Sapien

<400> 35  
 Met Ala Gln Lys Glu Gly Gly Arg Thr Val Pro Cys Cys Ser Arg Pro  
 1 5 10 15  
 Lys Val Ala Ala Leu Thr Ala Gly Thr Leu Leu Leu Leu Thr Ala Ile  
 20 25 30  
 Gly Ala Ala Ser Trp Ala Ile Val Ala Val Leu Leu Arg Ser Asp Gln  
 35 40 45  
 Glu Pro Leu Tyr Pro Val Gln Val Ser Ser Ala Asp Ala Arg Leu Met  
 50 55 60  
 Val Phe Asp Lys Thr Glu Gly Thr Trp Arg Leu Leu Cys Ser Ser Arg  
 65 70 75 80  
 Ser Asn Ala Arg Val Ala Gly Leu Ser Cys Glu Glu Met Gly Phe Leu  
 85 90 95  
 Arg Ala Leu Thr His Ser Glu Leu Asp Val Arg Thr Ala Gly Ala Asn  
 100 105 110  
 Gly Thr Ser Gly Phe Phe Cys Val Asp Glu Gly Arg Leu Pro His Thr  
 115 120 125  
 Gln Arg Leu Leu Glu Val Ile Ser Val Cys Asp Cys Pro Arg Gly Arg  
 130 135 140  
 Phe Leu Ala Ala Ile Cys Gln Asp Cys Gly Arg Arg Lys Leu Pro Val  
 145 150 155 160  
 Asp Arg Ile Val Gly Gly Arg Asp Thr Ser Leu Gly Arg Trp Pro Trp  
 165 170 175  
 Gln Val Ser Leu Arg Tyr Asp Gly Ala His Leu Cys Gly Gly Ser Leu  
 180 185 190  
 Leu Ser Gly Asp Trp Val Leu Thr Ala Ala His Cys Phe Pro Glu Arg  
 195 200 205  
 Asn Arg Val Leu Ser Arg Trp Arg Val Phe Ala Gly Ala Val Ala Gln  
 210 215 220  
 Ala Ser Pro His Gly Leu Gln Leu Gly Val Gln Ala Val Val Tyr His  
 225 230 235 240  
 Gly Gly Tyr Leu Pro Phe Arg Asp Pro Asn Ser Glu Glu Asn Ser Asn  
 245 250 255  
 Asp Ile Ala Leu Val His Leu Ser Ser Pro Leu Pro Leu Thr Glu Tyr  
 260 265 270  
 Ile Gln Pro Val Cys Leu Pro Ala Ala Gly Gln Ala Leu Val Asp Gly  
 275 280 285  
 Lys Ile Cys Thr Val Thr Gly Trp Gly Asn Thr Gln Tyr Tyr Gly Gln  
 290 295 300  
 Gln Ala Gly Val Leu Gln Glu Ala Arg Val Pro Ile Ile Ser Asn Asp  
 305 310 315 320  
 Val Cys Asn Gly Ala Asp Phe Tyr Gly Asn Gln Ile Lys Pro Lys Met

-73-

```

          325          330          335
Phe Cys Ala Gly Tyr Pro Glu Gly Gly Ile Asp Ala Cys Gln Gly Asp
          340          345          350
Ser Gly Gly Pro Phe Val Cys Glu Asp Ser Ile Ser Arg Thr Pro Arg
          355          360          365
Trp Arg Leu Cys Gly Ile Val Ser Trp Gly Thr Gly Cys Ala Leu Ala
          370          375          380
Gln Lys Pro Gly Val Tyr Thr Lys Val Ser Asp Phe Arg Glu Trp Ile
          385          390          395          400
Phe Gln Ala Ile Lys Thr His Ser Glu Ala Ser Gly Met Val Thr Gln
          405          410          415
Leu

```

&lt;210&gt; 36

&lt;211&gt; 2479

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (57)...(1535)

&lt;223&gt; Nucleotide sequence encoding human serine protease (TMPRSS2)

&lt;300&gt;

&lt;308&gt; GenBank U75329

&lt;309&gt; 1997-10-10

&lt;400&gt; 36

```

gtcatattga acattccaga tacctatcat tactcgatgc tgttgataac agcaag atg      59
                                     Met
                                     1

gct ttg aac tca ggg tca cca cca gct att gga cct tac tat gaa aac      107
Ala Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu Asn
          5          10          15

cat gga tac caa ccg gaa aac ccc tat ccc gca cag ccc act gtg gtc      155
His Gly Tyr Gln Pro Glu Asn Pro Tyr Pro Ala Gln Pro Thr Val Val
          20          25          30

ccc act gtc tac gag gtg cat ccg gct cag tac tac ccg tcc ccc gtg      203
Pro Thr Val Tyr Glu Val His Pro Ala Gln Tyr Tyr Pro Ser Pro Val
          35          40          45

ccc cag tac gcc ccg agg gtc ctg acg cag gct tcc aac ccc gtc gtc      251
Pro Gln Tyr Ala Pro Arg Val Leu Thr Gln Ala Ser Asn Pro Val Val
          50          55          60          65

tgc acg cag ccc aaa tcc cca tcc ggg aca gtg tgc acc tca aag act      299
Cys Thr Gln Pro Lys Ser Pro Ser Gly Thr Val Cys Thr Ser Lys Thr
          70          75          80

aag aaa gca ctg tgc atc acc ttg acc ctg ggg acc ttc ctc gtg gga      347
Lys Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val Gly
          85          90          95

gct gcg ctg gcc gct ggc cta ctc tgg aag ttc atg ggc agc aag tgc      395
Ala Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys Cys
          100          105          110

tcc aac tct ggg ata gag tgc gac tcc tca ggt acc tgc atc aac ccc      443

```

-74-

Ser	Asn	Ser	Gly	Ile	Glu	Cys	Asp	Ser	Ser	Gly	Thr	Cys	Ile	Asn	Pro		
115						120					125						
tct	aac	tgg	tgt	gat	ggc	gtg	tca	cac	tgc	ccc	ggc	ggg	gag	gac	gag	491	
Ser	Asn	Trp	Cys	Asp	Gly	Val	Ser	His	Cys	Pro	Gly	Gly	Glu	Asp	Glu		
130					135					140					145		
aat	cgg	tgt	gtt	cgc	ctc	tac	gga	cca	aac	ttc	atc	ctt	cag	atg	tac	539	
Asn	Arg	Cys	Val	Arg	Leu	Tyr	Gly	Pro	Asn	Phe	Ile	Leu	Gln	Met	Tyr		
				150					155					160			
tca	tct	cag	agg	aag	tcc	tgg	cac	cct	gtg	tgc	caa	gac	gac	tgg	aac	587	
Ser	Ser	Gln	Arg	Lys	Ser	Trp	His	Pro	Val	Cys	Gln	Asp	Asp	Trp	Asn		
				165				170					175				
gag	aac	tac	ggg	cgg	gcg	gcc	tgc	agg	gac	atg	ggc	tat	aag	aat	aat	635	
Glu	Asn	Tyr	Gly	Arg	Ala	Ala	Cys	Arg	Asp	Met	Gly	Tyr	Lys	Asn	Asn		
		180					185					190					
ttt	tac	tct	agc	caa	gga	ata	gtg	gat	gac	agc	gga	tcc	acc	agc	ttt	683	
Phe	Tyr	Ser	Ser	Gln	Gly	Ile	Val	Asp	Asp	Ser	Gly	Ser	Thr	Ser	Phe		
		195				200					205						
atg	aaa	ctg	aac	aca	agt	gcc	ggc	aat	gtc	gat	atc	tat	aaa	aaa	ctg	731	
Met	Lys	Leu	Asn	Thr	Ser	Ala	Gly	Asn	Val	Asp	Ile	Tyr	Lys	Lys	Leu		
210					215				220						225		
tac	cac	agt	gat	gcc	tgt	tct	tca	aaa	gca	gtg	gtt	tct	tta	cgc	tgt	779	
Tyr	His	Ser	Asp	Ala	Cys	Ser	Ser	Lys	Ala	Val	Val	Ser	Leu	Arg	Cys		
				230				235						240			
tta	gcc	tgc	ggg	gtc	aac	ttg	aac	tca	agc	cgc	cag	agc	agg	atc	gtg	827	
Leu	Ala	Cys	Gly	Val	Asn	Leu	Asn	Ser	Ser	Arg	Gln	Ser	Arg	Ile	Val		
			245					250					255				
ggc	ggt	gag	agc	gcg	ctc	ccg	ggg	gcc	tgg	ccc	tgg	cag	gtc	agc	ctg	875	
Gly	Gly	Glu	Ser	Ala	Leu	Pro	Gly	Ala	Trp	Pro	Trp	Gln	Val	Ser	Leu		
		260					265					270					
cac	gtc	cag	aac	gtc	cac	gtg	tgc	gga	ggc	tcc	atc	atc	acc	ccc	gag	923	
His	Val	Gln	Asn	Val	His	Val	Cys	Gly	Gly	Ser	Ile	Ile	Thr	Pro	Glu		
		275				280					285						
tgg	atc	gtg	aca	gcc	gcc	cac	tgc	gtg	gaa	aaa	cct	ctt	aac	aat	cca	971	
Trp	Ile	Val	Thr	Ala	Ala	His	Cys	Val	Glu	Lys	Pro	Leu	Asn	Asn	Pro		
					295					300					305		
tgg	cat	tgg	acg	gca	ttt	gcg	ggg	att	ttg	aga	caa	tct	ttc	atg	ttc	1019	
Trp	His	Trp	Thr	Ala	Phe	Ala	Gly	Ile	Leu	Arg	Gln	Ser	Phe	Met	Phe		
				310				315						320			
tat	gga	gcc	gga	tac	caa	gta	caa	aaa	gtg	att	tct	cat	cca	aat	tat	1067	
Tyr	Gly	Ala	Gly	Tyr	Gln	Val	Gln	Lys	Val	Ile	Ser	His	Pro	Asn	Tyr		
			325					330					335				
gac	tcc	aag	acc	aag	aac	aat	gac	att	gcg	ctg	atg	aag	ctg	cag	aag	1115	
Asp	Ser	Lys	Thr	Lys	Asn	Asn	Asp	Ile	Ala	Leu	Met	Lys	Leu	Gln	Lys		
		340					345					350					
cct	ctg	act	ttc	aac	gac	cta	gtg	aaa	cca	gtg	tgt	ctg	ccc	aac	cca	1163	
Pro	Leu	Thr	Phe	Asn	Asp	Leu	Val	Lys	Pro	Val	Cys	Leu	Pro	Asn	Pro		

-75-

355	360	365	
ggc atg atg ctg cag cca gaa cag ctc tgc tgg att tcc ggg tgg ggg			1211
Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp Gly			
370	375	380	385
gcc acc gag gag aaa ggg aag acc tca gaa gtg ctg aac gct gcc aag			1259
Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala Lys			
	390	395	400
gtg ctt ctc att gag aca cag aga tgc aac agc aga tat gtc tat gac			1307
Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr Asp			
	405	410	415
aac ctg atc aca cca gcc atg atc tgt gcc ggc ttc ctg cag ggg aac			1355
Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly Asn			
	420	425	430
gtc gat tct tgc cag ggt gac agt gga ggg cct ctg gtc act tgc aac			1403
Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser Asn			
	435	440	445
aac aat atc tgg tgg ctg ata ggg gat aca agc tgg ggt tct ggc tgt			1451
Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys			
	450	455	460
gcc aaa gct tac aga cca gga gtg tac ggg aat gtg atg gta ttc acg			1499
Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe Thr			
	470	475	480
gac tgg att tat cga caa atg aag gca aac ggc taa tccacatggt			1545
Asp Trp Ile Tyr Arg Gln Met Lys Ala Asn Gly *			
	485	490	
cttcgtcctt gacgtcggtt tacaagaaaa caatggggct ggttttgctt ccccggtgcat			1605
gatttactct tagagatgat tcagagggtca cttcattttt attaaacagt gaacttgtct			1665
ggcttttgga ctctctgccca tactgtgcag gctgcagtgg ctccccctgcc cagcctgctc			1725
tccctaaccct cttgtccgca aggggtgatg gccggctggt tgtgggcact ggcgggtcaat			1785
tgtggaagga agaggggttg aggtgcccc cattgagatc ttcttgctga gtcctttcca			1845
ggggccaatt ttggatgagc atggagctgt cacttctcag ctgctggatg acttgagatg			1905
aaaaaggaga gacatggaaa gggagacagc caggtggcac ctgcagcggc tgcctctggt			1965
ggccacttgg tagtgtcccc agcctacttc acaaggggat ttgtctgatg ggttcttaga			2025
gccttagcag ccctggatgg tggccagaaa taaagggacc agcccttcat ggggtggtgac			2085
gtggtagtca cttgtaaggg gaacagaaac atttttgttc ttatggggtg agaatataga			2145
cagtgcctct ggtgagaggg aagcaattga aaaggaactt gccctgagca ctctctggtgc			2205
aggtctccac ctgcacattg ggtggggctc ctgggaggga gactcagcct tctcctcat			2265
cctccctgac cctgctccta gcaccctgga gagtgaatgc cccttggtcc ctggcagggc			2325
gccaaagtttg gcaccatgtc ggctcttca ggcctgatag tcattggaaa ttgaggtcca			2385
tgggggaaat caaggatgct cagtttaagg tacactgttt ccatgttatg tttctacaca			2445
ttgatgggtg tgaccctgag ttcaaagcca tctt			2479

<210> 37  
 <211> 492  
 <212> PRT  
 <213> Homo sapien

<400> 37  
 Met Ala Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu  
 1 5 10 15  
 Asn His Gly Tyr Gln Pro Glu Asn Pro Tyr Pro Ala Gln Pro Thr Val  
 20 25 30

-76-

Val Pro Thr Val Tyr Glu Val His Pro Ala Gln Tyr Tyr Pro Ser Pro  
 35 40 45  
 Val Pro Gln Tyr Ala Pro Arg Val Leu Thr Gln Ala Ser Asn Pro Val  
 50 55 60  
 Val Cys Thr Gln Pro Lys Ser Pro Ser Gly Thr Val Cys Thr Ser Lys  
 65 70 75 80  
 Thr Lys Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val  
 85 90 95  
 Gly Ala Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys  
 100 105 110  
 Cys Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn  
 115 120 125  
 Pro Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp  
 130 135 140  
 Glu Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Met  
 145 150 155 160  
 Tyr Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp  
 165 170 175  
 Asn Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn  
 180 185 190  
 Asn Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser  
 195 200 205  
 Phe Met Lys Leu Asn Thr Ser Ala Gly Asn Val Asp Ile Tyr Lys Lys  
 210 215 220  
 Leu Tyr His Ser Asp Ala Cys Ser Ser Lys Ala Val Val Ser Leu Arg  
 225 230 235 240  
 Cys Leu Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile  
 245 250 255  
 Val Gly Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser  
 260 265 270  
 Leu His Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro  
 275 280 285  
 Glu Trp Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn  
 290 295 300  
 Pro Trp His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met  
 305 310 315 320  
 Phe Tyr Gly Ala Gly Tyr Gln Val Gln Lys Val Ile Ser His Pro Asn  
 325 330 335  
 Tyr Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln  
 340 345 350  
 Lys Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn  
 355 360 365  
 Pro Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp  
 370 375 380  
 Gly Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala  
 385 390 395 400  
 Lys Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr  
 405 410 415  
 Asp Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly  
 420 425 430  
 Asn Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser  
 435 440 445  
 Asn Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly  
 450 455 460  
 Cys Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe  
 465 470 475 480  
 Thr Asp Trp Ile Tyr Arg Gln Met Lys Ala Asn Gly  
 485 490

&lt;210&gt; 38

&lt;211&gt; 2079

-77-

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (251)...(1522)

<223> Nucleotide sequence encoding transmembrane  
protease, serine 4 (TMPRSS4)

&lt;300&gt;

&lt;308&gt; GenBank NM016425

&lt;309&gt; 2000-11-06

&lt;400&gt; 38

gagaggcagc agcttggttca gcggacaagg atgctgggcg tgagggacca aggcctgccc	60
tgcactcggg cctcctccag ccagtgtga ccagggactt ctgacctgct ggccagccag	120
gacctgtgtg gggaggccct cctgctgcct tggggtgaca atctcagctc caggctacag	180
ggagaccggg aggatcacag agccagcatg gtacaggatc ctgacagtga tcaacctctg	240
aacagcctcg atg tca aac ccc tgc gca aac ccc gta tcc cca tgg aga	289
Met Ser Asn Pro Cys Ala Asn Pro Val Ser Pro Trp Arg	
1 5 10	
cct tca gaa agt gtg ggg atc ccc atc atc ata gca cta ctg agc ctg	337
Pro Ser Glu Ser Val Gly Ile Pro Ile Ile Ile Ala Leu Leu Ser Leu	
15 20 25	
gcg agt atc atc att gtg gtt gtc ctc atc aag gtg att ctg gat aaa	385
Ala Ser Ile Ile Ile Val Val Val Leu Ile Lys Val Ile Leu Asp Lys	
30 35 40 45	
tac tac ttc ctc tgc ggg cag cct ctc cac ttc atc ccg agg aag cag	433
Tyr Tyr Phe Leu Cys Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln	
50 55 60	
ctg tgt gac gga gag ctg gac tgt ccc ttg ggg gag gac gag gag cac	481
Leu Cys Asp Gly Glu Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu His	
65 70 75	
tgt gtc aag agc ttc ccc gaa ggg cct gca gtg gca gtc cgc ctc tcc	529
Cys Val Lys Ser Phe Pro Glu Gly Pro Ala Val Ala Val Arg Leu Ser	
80 85 90	
aag gac cga tcc aca ctg cag gtg ctg gac tcg gcc aca ggg aac tgg	577
Lys Asp Arg Ser Thr Leu Gln Val Leu Asp Ser Ala Thr Gly Asn Trp	
95 100 105	
ttc tct gcc tgt ttc gac aac ttc aca gaa gct ctc gct gag aca gcc	625
Phe Ser Ala Cys Phe Asp Asn Phe Thr Glu Ala Leu Ala Glu Thr Ala	
110 115 120 125	
tgt agg cag atg ggc tac agc agc aaa ccc act ttc aga gct gtg gag	673
Cys Arg Gln Met Gly Tyr Ser Ser Lys Pro Thr Phe Arg Ala Val Glu	
130 135 140	
att ggc cca gac cag gat ctg gat gtt gtt gaa atc aca gaa aac agc	721
Ile Gly Pro Asp Gln Asp Leu Asp Val Val Glu Ile Thr Glu Asn Ser	
145 150 155	
cag gag ctt cgc atg cgg aac tca agt ggg ccc tgt ctc tca ggc tcc	769
Gln Glu Leu Arg Met Arg Asn Ser Ser Gly Pro Cys Leu Ser Gly Ser	
160 165 170	



-78-

ctg gtc tcc ctg cac tgt ctt gcc tgt ggg aag agc ctg aag acc ccc	817
Leu Val Ser Leu His Cys Leu Ala Cys Gly Lys Ser Leu Lys Thr Pro	
175 180 185	
cgt gtg gtg ggt ggg gag gag gcc tct gtg gat tct tgg cct tgg cag	865
Arg Val Val Gly Gly Glu Glu Ala Ser Val Asp Ser Trp Pro Trp Gln	
190 195 200 205	
gtc agc atc cag tac gac aaa cag cac gtc tgt gga ggg agc atc ctg	913
Val Ser Ile Gln Tyr Asp Lys Gln His Val Cys Gly Gly Ser Ile Leu	
210 215 220	
gac ccc cac tgg gtc ctc acg gca gcc cac tgc ttc agg aaa cat acc	961
Asp Pro His Trp Val Leu Thr Ala Ala His Cys Phe Arg Lys His Thr	
225 230 235	
gat gtg ttc aac tgg aag gtg cgg gca ggc tca gac aaa ctg ggc agc	1009
Asp Val Phe Asn Trp Lys Val Arg Ala Gly Ser Asp Lys Leu Gly Ser	
240 245 250	
ttc cca tcc ctg gct gtg gcc aag atc atc atc att gaa ttc aac ccc	1057
Phe Pro Ser Leu Ala Val Ala Lys Ile Ile Ile Ile Glu Phe Asn Pro	
255 260 265	
atg tac ccc aaa gac aat gac atc gcc ctc atg aag ctg cag ttc cca	1105
Met Tyr Pro Lys Asp Asn Asp Ile Ala Leu Met Lys Leu Gln Phe Pro	
270 275 280 285	
ctc act ttc tca ggc aca gtc agg ccc atc tgt ctg ccc ttc ttt gat	1153
Leu Thr Phe Ser Gly Thr Val Arg Pro Ile Cys Leu Pro Phe Phe Asp	
290 295 300	
gag gag ctc act cca gcc acc cca ctc tgg atc att gga tgg ggc ttt	1201
Glu Glu Leu Thr Pro Ala Thr Pro Leu Trp Ile Ile Gly Trp Gly Phe	
305 310 315	
acg aag cag aat gga ggg aag atg tct gac ata ctg ctg cag gcg tca	1249
Thr Lys Gln Asn Gly Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser	
320 325 330	
gtc cag gtc att gac agc aca cgg tgc aat gca gac gat gcg tac cag	1297
Val Gln Val Ile Asp Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln	
335 340 345	
ggg gaa gtc acc gag aag atg atg tgt gca ggc atc ccg gaa ggg ggt	1345
Gly Glu Val Thr Glu Lys Met Met Cys Ala Gly Ile Pro Glu Gly Gly	
350 355 360 365	
gtg gac acc tgc cag ggt gac agt ggt ggg ccc ctg atg tac caa tct	1393
Val Asp Thr Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser	
370 375 380	
gac cag tgg cat gtg gtg ggc atc gtt agc tgg ggc tat ggc tgc ggg	1441
Asp Gln Trp His Val Val Gly Ile Val Ser Trp Gly Tyr Gly Cys Gly	
385 390 395	
ggc ccg agc acc cca gga gta tac acc aag gtc tca gcc tat ctc aac	1489
Gly Pro Ser Thr Pro Gly Val Tyr Thr Lys Val Ser Ala Tyr Leu Asn	
400 405 410	

-79-

tgg atc tac aat gtc tgg aag gct gag ctg taa tgctgctgcc cctttgcagt 1542  
 Trp Ile Tyr Asn Val Trp Lys Ala Glu Leu \*  
 415 420

gctgggagcc gcttccttcc tgccctgccc acctggggat cccccaaagt cagacacaga 1602  
 gcaagagtcc ccttgggtac acccctctgc ccacagcctc agcatttctt ggagcagcaa 1662  
 agggcctcaa ttctgtgaag agaccctcgc agcccagagg cgcccagagg aagtcagcag 1722  
 ccctagctcg gccacacttg gtgctcccag catcccaggg agagacacag cccactgaac 1782  
 aaggtctcag ggggtattgct aagccaagaa ggaactttcc cacactactg aatggaagca 1842  
 ggctgtcttg taaaagccca gatcactgtg ggctggagag gagaaggaaa gggctctgcgc 1902  
 cagccctgtc cgtcttcacc catcccgaag cctactagag caagaaacca gttgtaatat 1962  
 aaaatgcact gccctactgt tggatatgact accgttacct actgttgtca ttgttattac 2022  
 agctatggcc actattatta aagagctgtg taacatcaaa aaaaaaaaaa aaaaaaa 2079

<210> 39  
 <211> 423  
 <212> PRT  
 <213> Homo sapien

<400> 39  
 Met Ser Asn Pro Cys Ala Asn Pro Val Ser Pro Trp Arg Pro Ser Glu  
 1 5 10 15  
 Ser Val Gly Ile Pro Ile Ile Ile Ala Leu Leu Ser Leu Ala Ser Ile  
 20 25 30  
 Ile Ile Val Val Val Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe  
 35 40 45  
 Leu Cys Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp  
 50 55 60  
 Gly Glu Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu His Cys Val Lys  
 65 70 75 80  
 Ser Phe Pro Glu Gly Pro Ala Val Ala Val Arg Leu Ser Lys Asp Arg  
 85 90 95  
 Ser Thr Leu Gln Val Leu Asp Ser Ala Thr Gly Asn Trp Phe Ser Ala  
 100 105 110  
 Cys Phe Asp Asn Phe Thr Glu Ala Leu Ala Glu Thr Ala Cys Arg Gln  
 115 120 125  
 Met Gly Tyr Ser Ser Lys Pro Thr Phe Arg Ala Val Glu Ile Gly Pro  
 130 135 140  
 Asp Gln Asp Leu Asp Val Val Glu Ile Thr Glu Asn Ser Gln Glu Leu  
 145 150 155 160  
 Arg Met Arg Asn Ser Ser Gly Pro Cys Leu Ser Gly Ser Leu Val Ser  
 165 170 175  
 Leu His Cys Leu Ala Cys Gly Lys Ser Leu Lys Thr Pro Arg Val Val  
 180 185 190  
 Gly Gly Glu Glu Ala Ser Val Asp Ser Trp Pro Trp Gln Val Ser Ile  
 195 200 205  
 Gln Tyr Asp Lys Gln His Val Cys Gly Gly Ser Ile Leu Asp Pro His  
 210 215 220  
 Trp Val Leu Thr Ala Ala His Cys Phe Arg Lys His Thr Asp Val Phe  
 225 230 235 240  
 Asn Trp Lys Val Arg Ala Gly Ser Asp Lys Leu Gly Ser Phe Pro Ser  
 245 250 255  
 Leu Ala Val Ala Lys Ile Ile Ile Ile Glu Phe Asn Pro Met Tyr Pro  
 260 265 270  
 Lys Asp Asn Asp Ile Ala Leu Met Lys Leu Gln Phe Pro Leu Thr Phe  
 275 280 285  
 Ser Gly Thr Val Arg Pro Ile Cys Leu Pro Phe Phe Asp Glu Glu Leu  
 290 295 300  
 Thr Pro Ala Thr Pro Leu Trp Ile Ile Gly Trp Gly Phe Thr Lys Gln  
 305 310 315 320  
 Asn Gly Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser Val Gln Val

-80-

```

          325          330          335
Ile Asp Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln Gly Glu Val
          340          345          350
Thr Glu Lys Met Met Cys Ala Gly Ile Pro Glu Gly Gly Val Asp Thr
          355          360          365
Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser Asp Gln Trp
          370          375          380
His Val Val Gly Ile Val Ser Trp Gly Tyr Gly Cys Gly Gly Pro Ser
          385          390          395          400
Thr Pro Gly Val Tyr Thr Lys Val Ser Ala Tyr Leu Asn Trp Ile Tyr
          405          410          415
Asn Val Trp Lys Ala Glu Leu
          420

```

<210> 40  
 <211> 1471  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> DESC1 gene

<221> misc feature  
 <222> (626)...(1324)  
 <223> protease domain

<221> CDS  
 <222> (56)...(1324)

```

<400> 40
tgacttggat gtagacctcg accttcacag gactcttcat tgctgggttg caatg atg      58
                                         Met
                                         1

tat cgg cca gat gtg gtg agg gct agg aaa aga gtt tgt tgg gaa ccc      106
Tyr Arg Pro Asp Val Val Arg Ala Arg Lys Arg Val Cys Trp Glu Pro
          5          10          15

tgg gtt atc ggc ctc gtc ats ttc ata tcc ctg att gtc ctg gca gtg      154
Trp Val Ile Gly Leu Val Xaa Phe Ile Ser Leu Ile Val Leu Ala Val
          20          25          30

tgc att gga stc act gtt cat tat gtg aga tat aat caa aag aag acc      202
Cys Ile Gly Xaa Thr Val His Tyr Val Arg Tyr Asn Gln Lys Lys Thr
          35          40          45

tac aat tac tat agc aca ttg tca ttt aca act gac aaa cta tat gct      250
Tyr Asn Tyr Tyr Ser Thr Leu Ser Phe Thr Thr Asp Lys Leu Tyr Ala
          50          55          60          65

gag ttt ggc aga gag gct tct aac aat ttt aca gaa atg agc cag aga      298
Glu Phe Gly Arg Glu Ala Ser Asn Asn Phe Thr Glu Met Ser Gln Arg
          70          75          80

ctt gaa tca atg gtg aaa aat gca ttt tat aaa tct cca tta agg gaa      346
Leu Glu Ser Met Val Lys Asn Ala Phe Tyr Lys Ser Pro Leu Arg Glu
          85          90          95

gaa ttt gtc aag tct cag gtt atc aag ttc agt caa cag aag cat gga      394
Glu Phe Val Lys Ser Gln Val Ile Lys Phe Ser Gln Gln Lys His Gly
          100          105          110

```

-81-

gtg ttg gct cat atg ctg ttg att tgt aga ttt cac tct act gag gat	442
Val Leu Ala His Met Leu Leu Ile Cys Arg Phe His Ser Thr Glu Asp	
115 120 125	
cct gaa act gta gat aaa att gtt caa ctt gtt tta cat gaa aag ctg	490
Pro Glu Thr Val Asp Lys Ile Val Gln Leu Val Leu His Glu Lys Leu	
130 135 140 145	
caa gat gct gta gga ccc cct aaa gta gat cct cac tca gtt aaa att	538
Gln Asp Ala Val Gly Pro Pro Lys Val Asp Pro His Ser Val Lys Ile	
150 155 160	
aaa aaa atc aac aag aca gaa aca gac agc tat cta aac cat tgc tgc	586
Lys Lys Ile Asn Lys Thr Glu Thr Asp Ser Tyr Leu Asn His Cys Cys	
165 170 175	
gga aca cga aga agt aaa act cta ggt cag agt ctc agg atc gtt ggt	634
Gly Thr Arg Arg Ser Lys Thr Leu Gly Gln Ser Leu Arg Ile Val Gly	
180 185 190	
ggg aca gaa gta gaa gag ggt gaa tgg ccc tgg cag gct agc ctg cag	682
Gly Thr Glu Val Glu Glu Gly Glu Trp Pro Trp Gln Ala Ser Leu Gln	
195 200 205	
tgg gat ggg agt cat cgc tgt gga gca acc tta att aat gcc aca tgg	730
Trp Asp Gly Ser His Arg Cys Gly Ala Thr Leu Ile Asn Ala Thr Trp	
210 215 220 225	
ctt gtg agt gct gct cac tgt ttt aca aca tat aag aac cct gcc aga	778
Leu Val Ser Ala Ala His Cys Phe Thr Thr Tyr Lys Asn Pro Ala Arg	
230 235 240	
tgg act gct tcc ttt gga gta aca ata aaa cct tcg aaa atg aaa cgg	826
Trp Thr Ala Ser Phe Gly Val Thr Ile Lys Pro Ser Lys Met Lys Arg	
245 250 255	
ggt ctc cgg aga ata att gtc cat gaa aaa tac aaa cac cca tca cat	874
Gly Leu Arg Arg Ile Ile Val His Glu Lys Tyr Lys His Pro Ser His	
260 265 270	
gac tat gat att tct ctt gca gag ctt tct agc cct gtt ccc tac aca	922
Asp Tyr Asp Ile Ser Leu Ala Glu Leu Ser Ser Pro Val Pro Tyr Thr	
275 280 285	
aat gca gta cat aga gtt tgt ctc cct gat gca tcc tat gag ttt caa	970
Asn Ala Val His Arg Val Cys Leu Pro Asp Ala Ser Tyr Glu Phe Gln	
290 295 300 305	
cca ggt gat gtg atg ttt gtg aca gga ttt gga gca ctg aaa aat gat	1018
Pro Gly Asp Val Met Phe Val Thr Gly Phe Gly Ala Leu Lys Asn Asp	
310 315 320	
ggt tac agt caa aat cat ctt cga caa gca cag gtg act ctc ata gac	1066
Gly Tyr Ser Gln Asn His Leu Arg Gln Ala Gln Val Thr Leu Ile Asp	
325 330 335	
gct aca act tgc aat gaa cct caa gct tac aat gac gcc ata act cct	1114
Ala Thr Thr Cys Asn Glu Pro Gln Ala Tyr Asn Asp Ala Ile Thr Pro	
340 345 350	

-82-

```

aga atg tta tgt gct ggc tcc tta gaa gga aaa aca gat gca tgc cag      1162
Arg Met Leu Cys Ala Gly Ser Leu Glu Gly Lys Thr Asp Ala Cys Gln
      355                      360                      365

ggg gac tct gga gga cca ctg gtt agt tca gat gct aga gat atc tgg      1210
Gly Asp Ser Gly Gly Pro Leu Val Ser Ser Asp Ala Arg Asp Ile Trp
      370                      375                      380                      385

tac ctt gct gga ata gtg agc tsg gga gat gaa tgt gcg aaa ccc aac      1258
Tyr Leu Ala Gly Ile Val Ser Xaa Gly Asp Glu Cys Ala Lys Pro Asn
                      390                      395                      400

aag cct ggt gtt tat act aga gtt acg gcc ttg cgg gac tgg att act      1306
Lys Pro Gly Val Tyr Thr Arg Val Thr Ala Leu Arg Asp Trp Ile Thr
                      405                      410                      415

tca aaa act ggt atc taa gagagaaaag cctcatggaa cagataacat      1354
Ser Lys Thr Gly Ile *
                      420

ttttttttgt tttttgggtg tggaggccat ttttagagat acagaattgg agaagacttg      1414
caaaacagct agatttgact gatctcaata aactgtttgc ttgatgcaaa aaaaaaa      1471

```

<210> 41  
 <211> 422  
 <212> PRT  
 <213> Homo Sapien

<220>  
 <221> VARIANT  
 <222> 24  
 <223> Xaa is Ile or Met

<221> VARIANT  
 <222> 37  
 <223> Xaa is Leu or Val

<221> VARIANT  
 <222> 393  
 <223> Xaa is Ser or Trp

```

<400> 41
Met Tyr Arg Pro Asp Val Val Arg Ala Arg Lys Arg Val Cys Trp Glu
  1          5          10          15
Pro Trp Val Ile Gly Leu Val Xaa Phe Ile Ser Leu Ile Val Leu Ala
      20          25          30
Val Cys Ile Gly Xaa Thr Val His Tyr Val Arg Tyr Asn Gln Lys Lys
      35          40          45
Thr Tyr Asn Tyr Tyr Ser Thr Leu Ser Phe Thr Thr Asp Lys Leu Tyr
      50          55          60
Ala Glu Phe Gly Arg Glu Ala Ser Asn Asn Phe Thr Glu Met Ser Gln
      65          70          75          80
Arg Leu Glu Ser Met Val Lys Asn Ala Phe Tyr Lys Ser Pro Leu Arg
      85          90          95
Glu Glu Phe Val Lys Ser Gln Val Ile Lys Phe Ser Gln Gln Lys His
      100          105          110
Gly Val Leu Ala His Met Leu Leu Ile Cys Arg Phe His Ser Thr Glu
      115          120          125
Asp Pro Glu Thr Val Asp Lys Ile Val Gln Leu Val Leu His Glu Lys
      130          135          140
Leu Gln Asp Ala Val Gly Pro Pro Lys Val Asp Pro His Ser Val Lys

```

-83-

145					150					155					160
Ile	Lys	Lys	Ile	Asn	Lys	Thr	Glu	Thr	Asp	Ser	Tyr	Leu	Asn	His	Cys
				165					170					175	
Cys	Gly	Thr	Arg	Arg	Ser	Lys	Thr	Leu	Gly	Gln	Ser	Leu	Arg	Ile	Val
			180					185					190		
Gly	Gly	Thr	Glu	Val	Glu	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Ala	Ser	Leu
		195					200					205			
Gln	Trp	Asp	Gly	Ser	His	Arg	Cys	Gly	Ala	Thr	Leu	Ile	Asn	Ala	Thr
	210					215					220				
Trp	Leu	Val	Ser	Ala	Ala	His	Cys	Phe	Thr	Thr	Tyr	Lys	Asn	Pro	Ala
	225			230					235						240
Arg	Trp	Thr	Ala	Ser	Phe	Gly	Val	Thr	Ile	Lys	Pro	Ser	Lys	Met	Lys
			245						250					255	
Arg	Gly	Leu	Arg	Arg	Ile	Ile	Val	His	Glu	Lys	Tyr	Lys	His	Pro	Ser
		260						265					270		
His	Asp	Tyr	Asp	Ile	Ser	Leu	Ala	Glu	Leu	Ser	Ser	Pro	Val	Pro	Tyr
	275						280					285			
Thr	Asn	Ala	Val	His	Arg	Val	Cys	Leu	Pro	Asp	Ala	Ser	Tyr	Glu	Phe
	290					295					300				
Gln	Pro	Gly	Asp	Val	Met	Phe	Val	Thr	Gly	Phe	Gly	Ala	Leu	Lys	Asn
	305				310					315					320
Asp	Gly	Tyr	Ser	Gln	Asn	His	Leu	Arg	Gln	Ala	Gln	Val	Thr	Leu	Ile
				325					330					335	
Asp	Ala	Thr	Thr	Cys	Asn	Glu	Pro	Gln	Ala	Tyr	Asn	Asp	Ala	Ile	Thr
		340						345					350		
Pro	Arg	Met	Leu	Cys	Ala	Gly	Ser	Leu	Glu	Gly	Lys	Thr	Asp	Ala	Cys
	355						360					365			
Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Ser	Ser	Asp	Ala	Arg	Asp	Ile
	370					375					380				
Trp	Tyr	Leu	Ala	Gly	Ile	Val	Ser	Xaa	Gly	Asp	Glu	Cys	Ala	Lys	Pro
	385				390					395					400
Asn	Lys	Pro	Gly	Val	Tyr	Thr	Arg	Val	Thr	Ala	Leu	Arg	Asp	Trp	Ile
			405						410					415	
Thr	Ser	Lys	Thr	Gly	Ile										
			420												

<210> 42  
 <211> 1257  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> CDS  
 <222> (1)...(1257)  
 <223> Nucleotide sequence encoding MTSP9

<400> 42	
atg atg tat cgg aca gta gga ttt ggc acc cga agc aga aat ctg aag	48
Met Met Tyr Arg Thr Val Gly Phe Gly Thr Arg Ser Arg Asn Leu Lys	
1 5 10 15	
cca tgg atg att gcc gtt ctc att gtg ttg tcc ctg aca gtg gtg gca	96
Pro Trp Met Ile Ala Val Leu Ile Val Leu Ser Leu Thr Val Val Ala	
20 25 30	
gtg acc ata ggt ctc ctg gtt cac ttc cta gta ttt gac caa aaa aag	144
Val Thr Ile Gly Leu Leu Val His Phe Leu Val Phe Asp Gln Lys Lys	
35 40 45	
gag tac tat cat ggc tcc ttt aaa att tta gat cca caa atc aat aac	192
Glu Tyr Tyr His Gly Ser Phe Lys Ile Leu Asp Pro Gln Ile Asn Asn	

-84-

50	55	60	
aat ttc gga caa agc aac aca tat caa ctt aag gac tta cga gag acg Asn Phe Gly Gln Ser Asn Thr Tyr Gln Leu Lys Asp Leu Arg Glu Thr 65 70 75 80			240
acc gaa aat ttg gtg gat gag ata ttt ata gat tca gcc tgg aag aaa Thr Glu Asn Leu Val Asp Glu Ile Phe Ile Asp Ser Ala Trp Lys Lys 85 90 95			288
aat tat atc aag aac caa gta gtc aga ctg act cca gag gaa gat ggt Asn Tyr Ile Lys Asn Gln Val Val Leu Thr Pro Glu Glu Asp Gly 100 105 110			336
gtg aaa gta gat gtc att atg gtg ttc cag ttc ccc tct act gaa caa Val Lys Val Asp Val Ile Met Val Phe Gln Phe Pro Ser Thr Glu Gln 115 120 125			384
agg gca gta aga gag aag aaa atc caa agc atc tta aat cag aag ata Arg Ala Val Arg Glu Lys Lys Ile Gln Ser Ile Leu Asn Gln Lys Ile 130 135 140			432
agg aat tta aga gcc ttg cca ata aat gcc tca tca gtt caa gtt aat Arg Asn Leu Arg Ala Leu Pro Ile Asn Ala Ser Ser Val Gln Val Asn 145 150 155 160			480
gca atg agc tca tca aca ggg gag tta act gtc caa gca agt tgt ggt Ala Met Ser Ser Ser Thr Gly Glu Leu Thr Val Gln Ala Ser Cys Gly 165 170 175			528
aaa cga gtt gtt cca tta aac gtc aac aga ata gca tct gga gtc att Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile 180 185 190			576
gca ccc aag gcg gcc tgg cct tgg caa gct tcc ctt cag tat gat aac Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn 195 200 205			624
atc cat cag tgt ggg gcc acc ttg att agt aac aca tgg ctt gtc act Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr 210 215 220			672
gca gca cac tgc ttc cag aag tat aaa aat cca cat caa tgg act gtt Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val 225 230 235 240			720
agt ttt gga aca aaa atc aac cct ccc tta atg aaa aga aat gtc aga Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg 245 250 255			768
aga ttt att atc cat gag aag tac cgc tct gca gca aga gag tac gac Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp 260 265 270			816
att gct gtt gtg cag gtc tct tcc aga gtc acc ttt tcg gat gac ata Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile 275 280 285			864
cgc cgg att tgt ttg cca gaa gcc tct gca tcc ttc caa cca aat ttg Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu 290 295 300			912

-85-

act gtc cac atc aca gga ttt gga gca ctt tac tat ggt ggg gaa tcc	960
Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser	
305 310 315 320	
caa aat gat ctc cga gaa gcc aga gtg aaa atc ata agt gac gat gtc	1008
Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val	
325 330 335	
tgc aag caa cca cag gtg tat ggc aat gat ata aaa cct gga atg ttc	1056
Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe	
340 345 350	
tgt gcc gga tat atg gaa gga att tat gat gcc tgc agg ggt gat tct	1104
Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser	
355 360 365	
ggg gga cct tta gtc aca agg gat ctg aaa gat acg tgg tat ctc att	1152
Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile	
370 375 380	
gga att gta agc tgg gga gat aac tgt ggt caa aag gac aag cct gga	1200
Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly	
385 390 395 400	
gtc tac aca caa gtg act tat tac cga aac tgg att gct tca aaa aca	1248
Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr	
405 410 415	
ggc atc taa	1257
Gly Ile *	

<210> 43  
 <211> 418  
 <212> PRT  
 <213> Homo sapien

<400> 43  
 Met Met Tyr Arg Thr Val Gly Phe Gly Thr Arg Ser Arg Asn Leu Lys  
 1 5 10 15  
 Pro Trp Met Ile Ala Val Leu Ile Val Leu Ser Leu Thr Val Val Ala  
 20 25 30  
 Val Thr Ile Gly Leu Leu Val His Phe Leu Val Phe Asp Gln Lys Lys  
 35 40 45  
 Glu Tyr Tyr His Gly Ser Phe Lys Ile Leu Asp Pro Gln Ile Asn Asn  
 50 55 60  
 Asn Phe Gly Gln Ser Asn Thr Tyr Gln Leu Lys Asp Leu Arg Glu Thr  
 65 70 75 80  
 Thr Glu Asn Leu Val Asp Glu Ile Phe Ile Asp Ser Ala Trp Lys Lys  
 85 90 95  
 Asn Tyr Ile Lys Asn Gln Val Val Arg Leu Thr Pro Glu Glu Asp Gly  
 100 105 110  
 Val Lys Val Asp Val Ile Met Val Phe Gln Phe Pro Ser Thr Glu Gln  
 115 120 125  
 Arg Ala Val Arg Glu Lys Lys Ile Gln Ser Ile Leu Asn Gln Lys Ile  
 130 135 140  
 Arg Asn Leu Arg Ala Leu Pro Ile Asn Ala Ser Ser Val Gln Val Asn  
 145 150 155 160  
 Ala Met Ser Ser Ser Thr Gly Glu Leu Thr Val Gln Ala Ser Cys Gly  
 165 170 175



-86-

Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile  
 180 185 190  
 Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn  
 195 200 205  
 Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr  
 210 215 220  
 Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val  
 225 230 235 240  
 Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg  
 245 250 255  
 Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp  
 260 265 270  
 Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile  
 275 280 285  
 Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu  
 290 295 300  
 Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser  
 305 310 315 320  
 Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val  
 325 330 335  
 Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe  
 340 345 350  
 Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser  
 355 360 365  
 Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile  
 370 375 380  
 Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly  
 385 390 395 400  
 Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr  
 405 410 415  
 Gly Ile

&lt;210&gt; 44

&lt;211&gt; 2130

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (0)...(2104)

 <223> Nucleotide sequence encoding MTSP10, including  
 MTSP10 protease domain

&lt;400&gt; 44

ata aac ctg gtt tat aca aca tct gcc ttc tcc aaa ttt tat gag cag	48
Ile Asn Leu Val Tyr Thr Thr Ser Ala Phe Ser Lys Phe Tyr Glu Gln	
1 5 10 15	
tct gtt gtt gca gat gtc agc agc aac aac aaa ggc ggc ctc ctt gtc	96
Ser Val Val Ala Asp Val Ser Ser Asn Asn Lys Gly Gly Leu Leu Val	
20 25 30	
cac ttt tgg att gtt ttt gtc atg cca cgt gcc aaa ggc cac atc ttc	144
His Phe Trp Ile Val Phe Val Met Pro Arg Ala Lys Gly His Ile Phe	
35 40 45	
tgt gaa gac tgt gtt gcc gcc atc ttg aag gac tcc atc cag aca agc	192
Cys Glu Asp Cys Val Ala Ala Ile Leu Lys Asp Ser Ile Gln Thr Ser	
50 55 60	
atc ata aac cgg acc tct gtg ggg agc ttg cag gga ctg gct gtg gac	240

-87-

Ile	Ile	Asn	Arg	Thr	Ser	Val	Gly	Ser	Leu	Gln	Gly	Leu	Ala	Val	Asp		
65					70					75					80		
atg	gac	tct	gtg	gta	cta	aat	gct	ggg	ctt	cgg	tca	gat	tac	tcg	tca	288	
Met	Asp	Ser	Val	Val	Leu	Asn	Ala	Gly	Leu	Arg	Ser	Asp	Tyr	Ser	Ser		
			85						90					95			
acc	ata	gga	tct	gac	aaa	ggc	tgc	tct	cag	tac	ttc	tat	gca	gag	cat	336	
Thr	Ile	Gly	Ser	Asp	Lys	Gly	Cys	Ser	Gln	Tyr	Phe	Tyr	Ala	Glu	His		
			100					105					110				
ctg	tct	ctc	cac	tac	ccg	ctg	gag	att	tct	gca	gcc	tca	ggg	agg	ctg	384	
Leu	Ser	Leu	His	Tyr	Pro	Leu	Glu	Ile	Ser	Ala	Ala	Ser	Gly	Arg	Leu		
			115				120					125					
atg	tgt	cac	ttc	aag	ctg	gtg	gcc	ata	gtg	ggc	tac	ctg	att	cgt	ctc	432	
Met	Cys	His	Phe	Lys	Leu	Val	Ala	Ile	Val	Gly	Tyr	Leu	Ile	Arg	Leu		
	130					135					140						
tca	atc	aag	tcc	atc	caa	atc	gaa	gcc	gac	aac	tgt	gtc	act	gac	tcc	480	
Ser	Ile	Lys	Ser	Ile	Gln	Ile	Glu	Ala	Asp	Asn	Cys	Val	Thr	Asp	Ser		
	145				150					155					160		
ctg	acc	att	tac	gac	tcc	ctt	ttg	ccc	atc	cgg	agc	agc	atc	ttg	tac	528	
Leu	Thr	Ile	Tyr	Asp	Ser	Leu	Leu	Pro	Ile	Arg	Ser	Ser	Ile	Leu	Tyr		
				165					170					175			
aga	att	tgt	gaa	ccc	aca	aga	aca	tta	atg	tca	ttt	gtt	tct	aca	aat	576	
Arg	Ile	Cys	Glu	Pro	Thr	Arg	Thr	Leu	Met	Ser	Phe	Val	Ser	Thr	Asn		
			180					185					190				
aat	ctc	atg	ttg	gtg	aca	ttt	aag	tct	cct	cat	ata	cgg	agg	ctc	tca	624	
Asn	Leu	Met	Leu	Val	Thr	Phe	Lys	Ser	Pro	His	Ile	Arg	Arg	Leu	Ser		
		195					200					205					
gga	atc	cgg	gca	tat	ttt	gag	gtc	att	cca	gaa	caa	aag	tgt	gaa	aac	672	
Gly	Ile	Arg	Ala	Tyr	Phe	Glu	Val	Ile	Pro	Glu	Gln	Lys	Cys	Glu	Asn		
	210					215					220						
aca	gtg	ttg	gtc	aaa	gac	atc	act	ggc	ttt	gaa	ggg	aaa	att	tca	agc	720	
Thr	Val	Leu	Val	Lys	Asp	Ile	Thr	Gly	Phe	Glu	Gly	Lys	Ile	Ser	Ser		
	225				230					235					240		
cca	tat	tac	ccg	agc	tac	tat	cct	cca	aaa	tgc	aag	tgt	acc	tgg	aaa	768	
Pro	Tyr	Tyr	Pro	Ser	Tyr	Tyr	Pro	Pro	Lys	Cys	Lys	Cys	Thr	Trp	Lys		
				245					250					255			
ttt	cag	act	tct	cta	tca	act	ctt	ggc	ata	gca	ctg	aaa	ttc	tat	aac	816	
Phe	Gln	Thr	Ser	Leu	Ser	Thr	Leu	Gly	Ile	Ala	Leu	Lys	Phe	Tyr	Asn		
			260					265					270				
tat	tca	ata	acc	aag	aag	agt	atg	aaa	ggc	tgt	gag	cat	gga	tgg	tgg	864	
Tyr	Ser	Ile	Thr	Lys	Lys	Ser	Met	Lys	Gly	Cys	Glu	His	Gly	Trp	Trp		
		275					280					285					
gaa	att	tat	gag	cac	atg	tac	tgt	ggc	tcc	tac	atg	gat	cat	cag	aca	912	
Glu	Ile	Tyr	Glu	His	Met	Tyr	Cys	Gly	Ser	Tyr	Met	Asp	His	Gln	Thr		
	290					295					300						
att	ttt	cga	gtg	ccc	agc	cct	ctg	gtt	cac	att	cag	ctc	cag	tgc	agt	960	
Ile	Phe	Arg	Val	Pro	Ser	Pro	Leu	Val	His	Ile	Gln	Leu	Gln	Cys	Ser		

-88-

305	310	315	320	
tca agg ctt tca ggc aag cca ctt ttg gca gaa tat ggc agt tac aac				1008
Ser Arg Leu Ser Gly Lys Pro Leu Leu Ala Glu Tyr Gly Ser Tyr Asn	325	330	335	
atc agt caa ccc tgc cct gtg gga tct ttt aga tgc tcc tcc ggt tta				1056
Ile Ser Gln Pro Cys Pro Val Gly Ser Phe Arg Cys Ser Ser Gly Leu	340	345	350	
tgt gtc cct cag gcc cag cgt ggt gat gga gta aat gac tgc ttt gat				1104
Cys Val Pro Gln Ala Gln Arg Gly Asp Gly Val Asn Asp Cys Phe Asp	355	360	365	
gaa agt gat gaa ctg ttt tgc gtg agc cct caa cct gcc tgc aat acc				1152
Glu Ser Asp Glu Leu Phe Cys Val Ser Pro Gln Pro Ala Cys Asn Thr	370	375	380	
agc tcc ttc agg cag cat ggc cct ctc atc tgt gat ggc ttc agg gac				1200
Ser Ser Phe Arg Gln His Gly Pro Leu Ile Cys Asp Gly Phe Arg Asp	385	390	395	400
tgt gag aat ggc cgg gat gag caa aac tgc act caa agt att cca tgc				1248
Cys Glu Asn Gly Arg Asp Glu Gln Asn Cys Thr Gln Ser Ile Pro Cys	405	410	415	
aac aac aga act ttt aag tgt ggc aat gat att tgc ttt agg aaa caa				1296
Asn Asn Arg Thr Phe Lys Cys Gly Asn Asp Ile Cys Phe Arg Lys Gln	420	425	430	
aat gca aaa tgt gat ggg aca gtg gat tgt cca gat gga agt gat gaa				1344
Asn Ala Lys Cys Asp Gly Thr Val Asp Cys Pro Asp Gly Ser Asp Glu	435	440	445	
gaa ggc tgc acc tgc agc agg agt tcc tcc gcc ctt cac cgc atc atc				1392
Glu Gly Cys Thr Cys Ser Arg Ser Ser Ser Ala Leu His Arg Ile Ile	450	455	460	
gga ggc aca gac acc ctg gag ggg ggt tgg ccg tgg cag gtc agc ctc				1440
Gly Gly Thr Asp Thr Leu Glu Gly Gly Trp Pro Trp Gln Val Ser Leu	465	470	475	480
cac ttt gtt gga tct gcc tac tgt ggt gcc tca gtc atc tcc agg gag				1488
His Phe Val Gly Ser Ala Tyr Cys Gly Ala Ser Val Ile Ser Arg Glu	485	490	495	
tgg ctt ctt tct gca gcc cac tgt ttt cat gga aac agg ctg tca gat				1536
Trp Leu Leu Ser Ala Ala His Cys Phe His Gly Asn Arg Leu Ser Asp	500	505	510	
ccc aca cca tgg act gca cac ctc ggg atg tat gtt cag ggg aat gcc				1584
Pro Thr Pro Trp Thr Ala His Leu Gly Met Tyr Val Gln Gly Asn Ala	515	520	525	
aag ttt gtc tcc ccg gtg aga aga att gtg gtc cac gag tac tat aac				1632
Lys Phe Val Ser Pro Val Arg Arg Ile Val Val His Glu Tyr Tyr Asn	530	535	540	
agt cag act ttt gat tat gat att gct ttg cta cag ctc agt att gcc				1680
Ser Gln Thr Phe Asp Tyr Asp Ile Ala Leu Leu Gln Leu Ser Ile Ala	545	550	555	560

-89-

tgg cct gag acc ctg aaa cag ctc att cag cca ata tgc att cct ccc	1728
Trp Pro Glu Thr Leu Lys Gln Leu Ile Gln Pro Ile Cys Ile Pro Pro	
565 570 575	
act ggt cag aga gtt cgc agt ggg gag aag tgc tgg gta act ggc tgg	1776
Thr Gly Gln Arg Val Arg Ser Gly Glu Lys Cys Trp Val Thr Gly Trp	
580 585 590	
ggg cga aga cac gaa gca gat aat aaa ggc tcc ctc gtt ctg cag caa	1824
Gly Arg Arg His Glu Ala Asp Asn Lys Gly Ser Leu Val Leu Gln Gln	
595 600 605	
gcg gag gta gag ctc att gat caa acg ctc tgt gtt tcc acc tac ggg	1872
Ala Glu Val Glu Leu Ile Asp Gln Thr Leu Cys Val Ser Thr Tyr Gly	
610 615 620	
atc atc act tct cgg atg ctc tgt gca ggc ata atg tca ggc aag aga	1920
Ile Ile Thr Ser Arg Met Leu Cys Ala Gly Ile Met Ser Gly Lys Arg	
625 630 635 640	
gat gcc tgc aaa gga gat tgc ggt gga cct tta tct tgt cga aga aaa	1968
Asp Ala Cys Lys Gly Asp Ser Gly Gly Pro Leu Ser Cys Arg Arg Lys	
645 650 655	
agt gat gga aaa tgg att ttg act ggc att gtt agc tgg gga cat gga	2016
Ser Asp Gly Lys Trp Ile Leu Thr Gly Ile Val Ser Trp Gly His Gly	
660 665 670	
tgt gga cga cca aac ttt cct ggt gtt tac aca agg gtg tca aac ttt	2064
Cys Gly Arg Pro Asn Phe Pro Gly Val Tyr Thr Arg Val Ser Asn Phe	
675 680 685	
gtt ccc tgg att cat aaa tat gtc cct tct ctt ttg taa t tgcaaaaaaa	2114
Val Pro Trp Ile His Lys Tyr Val Pro Ser Leu Leu *	
690 695 700	
aaaaaaaaa aaaaaa	2130

&lt;210&gt; 45

&lt;211&gt; 700

&lt;212&gt; PRT

&lt;213&gt; Homo Sapien

&lt;400&gt; 45

Ile Asn Leu Val Tyr Thr Thr Ser Ala Phe Ser Lys Phe Tyr Glu Gln	
1 5 10 15	
Ser Val Val Ala Asp Val Ser Ser Asn Asn Lys Gly Gly Leu Leu Val	
20 25 30	
His Phe Trp Ile Val Phe Val Met Pro Arg Ala Lys Gly His Ile Phe	
35 40 45	
Cys Glu Asp Cys Val Ala Ala Ile Leu Lys Asp Ser Ile Gln Thr Ser	
50 55 60	
Ile Ile Asn Arg Thr Ser Val Gly Ser Leu Gln Gly Leu Ala Val Asp	
65 70 75 80	
Met Asp Ser Val Val Leu Asn Ala Gly Leu Arg Ser Asp Tyr Ser Ser	
85 90 95	
Thr Ile Gly Ser Asp Lys Gly Cys Ser Gln Tyr Phe Tyr Ala Glu His	
100 105 110	
Leu Ser Leu His Tyr Pro Leu Glu Ile Ser Ala Ala Ser Gly Arg Leu	
115 120 125	

-90-

Met	Cys	His	Phe	Lys	Leu	Val	Ala	Ile	Val	Gly	Tyr	Leu	Ile	Arg	Leu
	130					135					140				
Ser	Ile	Lys	Ser	Ile	Gln	Ile	Glu	Ala	Asp	Asn	Cys	Val	Thr	Asp	Ser
145					150					155					160
Leu	Thr	Ile	Tyr	Asp	Ser	Leu	Leu	Pro	Ile	Arg	Ser	Ser	Ile	Leu	Tyr
				165					170					175	
Arg	Ile	Cys	Glu	Pro	Thr	Arg	Thr	Leu	Met	Ser	Phe	Val	Ser	Thr	Asn
			180					185					190		
Asn	Leu	Met	Leu	Val	Thr	Phe	Lys	Ser	Pro	His	Ile	Arg	Arg	Leu	Ser
	195						200					205			
Gly	Ile	Arg	Ala	Tyr	Phe	Glu	Val	Ile	Pro	Glu	Gln	Lys	Cys	Glu	Asn
	210					215					220				
Thr	Val	Leu	Val	Lys	Asp	Ile	Thr	Gly	Phe	Glu	Gly	Lys	Ile	Ser	Ser
225					230					235					240
Pro	Tyr	Tyr	Pro	Ser	Tyr	Tyr	Pro	Pro	Lys	Cys	Lys	Cys	Thr	Trp	Lys
				245					250					255	
Phe	Gln	Thr	Ser	Leu	Ser	Thr	Leu	Gly	Ile	Ala	Leu	Lys	Phe	Tyr	Asn
			260					265					270		
Tyr	Ser	Ile	Thr	Lys	Lys	Ser	Met	Lys	Gly	Cys	Glu	His	Gly	Trp	Trp
	275						280					285			
Glu	Ile	Tyr	Glu	His	Met	Tyr	Cys	Gly	Ser	Tyr	Met	Asp	His	Gln	Thr
	290					295					300				
Ile	Phe	Arg	Val	Pro	Ser	Pro	Leu	Val	His	Ile	Gln	Leu	Gln	Cys	Ser
305					310					315					320
Ser	Arg	Leu	Ser	Gly	Lys	Pro	Leu	Leu	Ala	Glu	Tyr	Gly	Ser	Tyr	Asn
				325					330					335	
Ile	Ser	Gln	Pro	Cys	Pro	Val	Gly	Ser	Phe	Arg	Cys	Ser	Ser	Gly	Leu
			340					345						350	
Cys	Val	Pro	Gln	Ala	Gln	Arg	Gly	Asp	Gly	Val	Asn	Asp	Cys	Phe	Asp
		355					360					365			
Glu	Ser	Asp	Glu	Leu	Phe	Cys	Val	Ser	Pro	Gln	Pro	Ala	Cys	Asn	Thr
	370					375					380				
Ser	Ser	Phe	Arg	Gln	His	Gly	Pro	Leu	Ile	Cys	Asp	Gly	Phe	Arg	Asp
385					390					395					400
Cys	Glu	Asn	Gly	Arg	Asp	Glu	Gln	Asn	Cys	Thr	Gln	Ser	Ile	Pro	Cys
				405					410					415	
Asn	Asn	Arg	Thr	Phe	Lys	Cys	Gly	Asn	Asp	Ile	Cys	Phe	Arg	Lys	Gln
			420					425					430		
Asn	Ala	Lys	Cys	Asp	Gly	Thr	Val	Asp	Cys	Pro	Asp	Gly	Ser	Asp	Glu
		435					440					445			
Glu	Gly	Cys	Thr	Cys	Ser	Arg	Ser	Ser	Ser	Ala	Leu	His	Arg	Ile	Ile
	450					455					460				
Gly	Gly	Thr	Asp	Thr	Leu	Glu	Gly	Gly	Trp	Pro	Trp	Gln	Val	Ser	Leu
465					470					475					480
His	Phe	Val	Gly	Ser	Ala	Tyr	Cys	Gly	Ala	Ser	Val	Ile	Ser	Arg	Glu
				485					490					495	
Trp	Leu	Leu	Ser	Ala	Ala	His	Cys	Phe	His	Gly	Asn	Arg	Leu	Ser	Asp
			500					505					510		
Pro	Thr	Pro	Trp	Thr	Ala	His	Leu	Gly	Met	Tyr	Val	Gln	Gly	Asn	Ala
		515					520					525			
Lys	Phe	Val	Ser	Pro	Val	Arg	Arg	Ile	Val	Val	His	Glu	Tyr	Tyr	Asn
	530						535				540				
Ser	Gln	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Gln	Leu	Ser	Ile	Ala
545					550					555					560
Trp	Pro	Glu	Thr	Leu	Lys	Gln	Leu	Ile	Gln	Pro	Ile	Cys	Ile	Pro	Pro
				565					570					575	
Thr	Gly	Gln	Arg	Val	Arg	Ser	Gly	Glu	Lys	Cys	Trp	Val	Thr	Gly	Trp
			580					585					590		
Gly	Arg	Arg	His	Glu	Ala	Asp	Asn	Lys	Gly	Ser	Leu	Val	Leu	Gln	Gln
		595					600					605			
Ala	Glu	Val	Glu	Leu	Ile	Asp	Gln	Thr	Leu	Cys	Val	Ser	Thr	Tyr	Gly

-91-

610		615		620											
Ile	Ile	Thr	Ser	Arg	Met	Leu	Cys	Ala	Gly	Ile	Met	Ser	Gly	Lys	Arg
625					630					635					640
Asp	Ala	Cys	Lys	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ser	Cys	Arg	Arg	Lys
				645					650					655	
Ser	Asp	Gly	Lys	Trp	Ile	Leu	Thr	Gly	Ile	Val	Ser	Trp	Gly	His	Gly
			660					665					670		
Cys	Gly	Arg	Pro	Asn	Phe	Pro	Gly	Val	Tyr	Thr	Arg	Val	Ser	Asn	Phe
		675					680					685			
Val	Pro	Trp	Ile	His	Lys	Tyr	Val	Pro	Ser	Leu	Leu				
	690					695					700				

&lt;210&gt; 46

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 46

Leu	Arg	Ala	Xaa	Gly	Arg	Ala	Xaa
1				5			

&lt;210&gt; 47

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Quat: (R)-Glu(alpha-(3-amidinobenzyl))

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Xaa is Alanine-therapeutic agent

&lt;400&gt; 47

Leu	Arg	Ala	Xaa	Ala	Arg	Ala	Xaa
1				5			

&lt;210&gt; 48

&lt;211&gt; 8

&lt;212&gt; PRT

-92-

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 48

Leu Arg Ser Xaa Gly Arg Ala Xaa  
1 5

&lt;210&gt; 49

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 49

Leu Arg Ser Xaa Ala Arg Ala Xaa  
1 5

&lt;210&gt; 50

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 50

Leu Arg Pro Arg Phe Lys Ile Xaa  
1 5

-93-

<210> 51  
<211> 7  
<212> PRT  
<213> Artificial sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7  
<223> Isoleucine-therapeutic agent

<400> 51  
Arg Pro Arg Phe Lys Ile Xaa  
1 5

<210> 52  
<211> 6  
<212> PRT  
<213> Artificial sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Isoleucine-therapeutic agent

<400> 52  
Pro Arg Phe Lys Ile Xaa  
1 5

<210> 53  
<211> 8  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 8  
<223> Alanine-therapeutic agent

<400> 53  
Leu Arg Ser Lys Ser Arg Ala Xaa  
1 5

<210> 54  
<211> 7  
<212> PRT  
<213> Artificial sequence



-94-

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent

<400> 54  
Arg Ser Lys Ser Arg Ala Xaa  
1 5

<210> 55  
<211> 6  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 55  
Ser Lys Ser Arg Ala Xaa  
1 5

<210> 56  
<211> 8  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Isoleucine-therapeutic agent

<400> 56  
Leu Arg Pro Arg Phe Arg Ile Xaa  
1 5

<210> 57  
<211> 7  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION

-95-

<222> 1  
<221> MOD\_RES  
<222> 7  
<223> Isoleucine-therapeutic agent  
  
<400> 57  
Arg Pro Arg Phe Arg Ile Xaa  
1 5  
  
<210> 58  
<211> 6  
<212> PRT  
<213> Artificial sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Isoleucine-therapeutic agent  
  
<400> 58  
Pro Arg Phe Arg Ile Xaa  
1 5  
  
<210> 59  
<211> 8  
<212> PRT  
<213> Artificial sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 8  
<223> Isoleucine-therapeutic agent  
  
<400> 59  
Leu Arg Ser Arg Ser Arg Ala Xaa  
1 5  
  
<210> 60  
<211> 7  
<212> PRT  
<213> Artificial sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7

-96-

<223> Alanine-therapeutic agent

<400> 60

Arg Ser Arg Ser Arg Ala Xaa  
1 5

<210> 61

<211> 6

<212> PRT

<213> Artificial sequence

<220>

<223> Conjugate

<221> ACETYLATION

<222> 1

<221> MOD\_RES

<222> 6

<223> Alanine-therapeutic agent

<400> 61

Ser Arg Ser Arg Ala Xaa  
1 5

<210> 62

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<223> Conjugate

<221> ACETYLATION

<222> 1

<221> MOD\_RES

<222> 4

<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

<222> 8

<223> Alanine-therapeutic agent

<400> 62

Ieu Arg Ala Xaa Gly Arg Ala Xaa  
1 5

<210> 63

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<223> Conjugate

<221> ACETYLATION

<222> 1

<221> MOD\_RES

<222> 4

-97-

<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

<222> 8

<223> Alanine-therapeutic agent

<400> 63

Leu Arg Ala Xaa Ala Arg Ala Xaa  
1 5

<210> 64

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<223> Conjugate

<221> ACETYLTATION

<222> 1

<221> MOD\_RES

<222> 4

<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

<222> 8

<223> Alanine-therapeutic agent

<400> 64

Leu Arg Ser Xaa Gly Arg Ala Xaa  
1 5

<210> 65

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<223> Conjugate

<221> ACETYLTATION

<222> 1

<221> MOD\_RES

<222> 4

<223> Xaa is Quat:  
(R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

<222> 8

<223> Alanine-therapeutic agent

<400> 65

Leu Arg Ser Xaa Ala Arg Ala Xaa  
1 5

<210> 66

<211> 8

<212> PRT

<213> Artificial sequence

-98-

<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 8  
<223> Isoleucine-therapeutic agent

<400> 66  
Leu Arg Pro Arg Phe Lys Ile Xaa  
1 5

<210> 67  
<211> 7  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7  
<223> Isoleucine-therapeutic agent

<400> 67  
Arg Pro Arg Phe Lys Ile Xaa  
1 5

<210> 68  
<211> 6  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Isoleucine-therapeutic agent

<400> 68  
Pro Arg Phe Lys Ile Xaa  
1 5

<210> 69  
<211> 8  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate  
  
<221> ACETYLATION

-99-

<222> 1  
<221> MOD\_RES  
<222> 8  
<223> Alanine-therapeutic agent  
  
<400> 69  
Leu Arg Ser Lys Ser Arg Ala Xaa  
1 5  
  
<210> 70  
<211> 7  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent  
  
<400> 70  
Arg Ser Lys Ser Arg Ala Xaa  
1 5  
  
<210> 71  
<211> 6  
<212> PRT  
<213> Artifical sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> (1)...(0)  
  
<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent  
  
<400> 71  
Ser Lys Ser Arg Ala Xaa  
1 5  
  
<210> 72  
<211> 8  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 8

-100-

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 72

Leu Arg Pro Arg Phe Arg Ile Xaa  
1 5

&lt;210&gt; 73

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 73

Arg Pro Arg Phe Arg Ile Xaa  
1 5

&lt;210&gt; 74

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 74

Pro Arg Phe Arg Ile Xaa  
1 5

&lt;210&gt; 75

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-Therapeutic Agent

&lt;400&gt; 75

Leu Arg Ser Arg Ser Arg Ala Xaa

-101-

1 5

<210> 76  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent

<400> 76  
Arg Ser Arg Ser Arg Ala Xaa  
1 5

<210> 77  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 77  
Ser Arg Ser Arg Ala Xaa  
1 5

<210> 78  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 78  
Xaa Pro Arg Ala Xaa  
1 5

<210> 79



-102-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-HHT:  
HHT is hexahydrotyrosol

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 79  
Xaa Gly Arg Ala Xaa  
1 5

<210> 80  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-p-tosyl-Gly

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 80  
Xaa Pro Arg Ala Xaa  
1 5

<210> 81  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Benzoyl-Val

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 81  
Xaa Gly Arg Ala Xaa  
1 5

<210> 82

-103-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH3SO2-D-HHT:  
HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 82  
Xaa Gly Arg Ala Xaa  
1 5

<210> 83  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-alpha-Z-D-Arg:  
Z is benzyloxycarbonyl

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 83  
Xaa Gly Arg Ala Xaa  
1 5

<210> 84  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 84  
Xaa Gly Arg Ala Xaa  
1 5

-104-

<210> 85  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Ile

<221> MOD\_RES  
<222> 5  
<223> Alanine- therapeutic agent

<400> 85  
Xaa Pro Arg Ala Xaa  
1 5

<210> 86  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Cbo-L-(gamma)Glu(alpha-t-BuO):  
Cbo is carbobenzoxy

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 86  
Xaa Arg Ala Ala Xaa  
1 5

<210> 87  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Pro

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 87  
Xaa Phe Arg Ala Xaa  
1 5

-105-

<210> 88  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MUTAGEN  
<222> 5  
<223> Alanine-therapeutic agent

<400> 88  
Xaa Leu Arg Ala Xaa  
1 5

<210> 89  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile:  
Bz is benzoyl

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 89  
Xaa Glu Gly Arg Ala Xaa  
1 5

<210> 90  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile:  
Bz is benzoyl

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 90  
Xaa Xaa Gly Arg Ala Xaa  
1 5

-106-

<210> 91  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Benzoyl-Pro

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 91  
Xaa Phe Arg Ala Xaa  
1 5

<210> 92  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Phe

<221> MOD\_RES  
<222> 2  
<223> pipecolinic acid

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 92  
Xaa Xaa Arg Ala Xaa  
1 5

<210> 93  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 93

-107-

Xaa Leu Lys Ala Xaa  
1 5

<210> 94  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Nle

<221> MOD\_RES  
<222> 2  
<223> HHT:  
HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 94  
Xaa Xaa Lys Ala Xaa  
1 5

<210> 95  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent

<400> 95  
Xaa Arg Thr Lys Arg Ala Xaa  
1 5

<210> 96  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-Arg

<221> MOD\_RES

-108-

<222> 6  
<223> Alanine-therapeutic agent

<400> 96  
Xaa Gln Arg Arg Ala Xaa  
1 5

<210> 97  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Boc-Gln:  
Boc is t-butoxycarbonyl

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 97  
Xaa Gly Arg Ala Xaa  
1 5

<210> 98  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Z-Arg:  
Z is benzyloxycarbonyl

<221> MOD\_RES  
<222> 4  
<223> Alanine-therapeutic agent

<400> 98  
Xaa Arg Ala Xaa  
1

<210> 99  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-HHT: HHT is hexahydrotyrosol

-109-

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 99  
Xaa Ala Arg Ala Xaa  
1 5

<210> 100  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-CHT:  
HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 100  
Xaa Gly Arg Ala Xaa  
1 5

<210> 101  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is MeSO<sub>2</sub>-D-Phe

<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 101  
Xaa Pro Arg Ala Xaa  
1 5

<210> 102  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1



-110-

<223> Xaa is delta-Z-D-Lys: Z is benzyloxycarbonyl

<221> MOD\_RES

<222> 5

<223> Alanine-therapeutic agent

<400> 102

Xaa Pro Arg Ala Xaa

1 5

<210> 103

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Conjugate

<221> MOD\_RES

<222> 1

<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-CHA:

CHA is cyclohexylalanyl

<221> MOD\_RES

<222> 2

<223> Xaa is But-Arg

<221> MOD\_RES

<222> 4

<223> Alanine-therapeutic agent

<400> 103

Xaa Xaa Ala Xaa

1

<210> 104

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Conjugate

<221> ACETYLATION

<222> 1

<221> MOD\_RES

<222> 6

<223> Alanine-therapeutic agent

<400> 104

Arg Gln Ser Arg Ala Xaa

1 5

<210> 105

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Conjugate

-111-

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Ala-therapeutic agent

&lt;400&gt; 105

Arg Arg Gln Ser Arg Ala Xaa  
1 5

&lt;210&gt; 106

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 106

Leu Arg Arg Gln Ser Arg Ala Xaa  
1 5

&lt;210&gt; 107

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 107

Arg Gln Ser Arg Xaa  
1 5

&lt;210&gt; 108

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

-112-

<222> 6  
<223> Alanine-therapeutic agent

<400> 108  
Arg Arg Gln Ser Arg Xaa  
1 5

<210> 109  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Glycine-therapeutic agent

<400> 109  
Leu Arg Arg Gln Ser Arg Gly Xaa  
1 5

<210> 110  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent

<400> 110  
Leu Arg Arg Gln Ser Arg Xaa  
1 5

<210> 111  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 6  
<223> Isoleucine-therapeutic agent

<400> 111

-113-

Arg Arg Gln Ser Arg Xaa  
 1 5

<210> 112  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conjugate

<221> ACETYLATION  
 <222> 1

<221> MOD\_RES  
 <222> 8  
 <223> Isoleucine-therapeutic agent

<400> 112  
 Leu Arg Arg Gln Ser Arg Ala Xaa  
 1 5

<210> 113  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conjugate

<221> ACETYLATION  
 <222> (1)...(0)

<221> MOD\_RES  
 <222> 4  
 <223> Xaa is Quat: (R)-Glu(Alpha-(3-amidinobenzyl))

<221> MOD\_RES  
 <222> 8  
 <223> Leucine-therapeutic agent

<400> 113  
 Leu Arg Ala Xaa Gly Arg Ser Xaa  
 1 5

<210> 114  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conjugate

<221> ACETYLATION  
 <222> 1

<221> MOD\_RES  
 <222> 4  
 <223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

-114-

<222> 8  
<223> Leucine-therapeutic agent

<400> 114  
Leu Arg Ala Xaa Ala Arg Ser Xaa  
1 5

<210> 115  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 115  
Leu Arg Ser Xaa Gly Arg Ser Xaa  
1 5

<210> 116  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 116  
Leu Arg Ser Xaa Ala Arg Ser Xaa  
1 5

<210> 117  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

-115-

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 117

Leu Arg Pro Arg Phe Lys Ser Xaa  
1 5

&lt;210&gt; 118

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 118

Arg Pro Arg Phe Lys Ser Xaa  
1 5

&lt;210&gt; 119

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 119

Pro Arg Phe Lys Ser Xaa  
1 5

&lt;210&gt; 120

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

-116-

<222> 8  
<223> Leucine-therapeutic agent

<400> 120  
Leu Arg Ser Lys Ser Arg Ser Xaa  
1 5

<210> 121  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 121  
Arg Ser Lys Ser Arg Ser Xaa  
1 5

<210> 122  
<211> 6  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 122  
Ser Lys Ser Arg Ser Xaa  
1 5

<210> 123  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 123

-117-

Leu Arg Pro Arg Phe Arg Ser Xaa  
1 5

<210> 124  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 124  
Arg Pro Arg Phe Arg Ser Xaa  
1 5

<210> 125  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 125  
Pro Arg Phe Arg Ser Xaa  
1 5

<210> 126  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 126  
Leu Arg Ser Arg Ser Arg Ser Xaa  
1 5

<210> 127



-118-

<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 127  
Arg Ser Arg Ser Arg Ser Xaa  
1 5

<210> 128  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 128  
Ser Arg Ser Arg Ser Xaa  
1 5

<210> 129  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 129  
Leu Arg Ala Xaa Gly Arg Ser Xaa  
1 5

<210> 130

-119-

<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 130  
Leu Arg Ala Xaa Ala Arg Ser Xaa  
1 5

<210> 131  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 131  
Leu Arg Ser Xaa Gly Arg Ser Xaa  
1 5

<210> 132  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES

-120-

<222> 8  
<223> Leucine-therapeutic agent

<400> 132  
Leu Arg Ser Xaa Ala Arg Ser Xaa  
1 5

<210> 133  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 133  
Leu Arg Pro Arg Phe Lys Ser Xaa  
1 5

<210> 134  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 134  
Arg Pro Arg Phe Lys Ser Xaa  
1 5

<210> 135  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 135

-121-

Pro Arg Phe Lys Ser Xaa  
1 5

<210> 136  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 136  
Leu Arg Ser Lys Ser Arg Ser Xaa  
1 5

<210> 137  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 137  
Arg Ser Lys Ser Arg Ser Xaa  
1 5

<210> 138  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 138  
Ser Lys Ser Arg Ser Xaa  
1 5

<210> 139

-122-

<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 139  
Leu Arg Pro Arg Phe Arg Ser Xaa  
1 5

<210> 140  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 140  
Arg Pro Arg Phe Arg Ser Xaa  
1 5

<210> 141  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 141  
Pro Arg Phe Arg Ser Xaa  
1 5

<210> 142  
<211> 8  
<212> PRT  
<213> Artificial Sequence

-123-

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 142  
Leu Arg Ser Arg Ser Arg Ser Xaa  
1 5

<210> 143  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 143  
Arg Ser Arg Ser Arg Ser Xaa  
1 5

<210> 144  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 144  
Ser Arg Ser Arg Ser Xaa  
1 5

<210> 145  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES

-124-

<222> 1  
<223> Xaa is pyroglutamic acid  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 145  
Xaa Pro Arg Ser Xaa  
1 5

<210> 146  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-HHT;  
HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 146  
Xaa Gly Arg Ser Xaa  
1 5

<210> 147  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is n-p-tosyl-Gly

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 147  
Xaa Pro Arg Ser Xaa  
1 5

<210> 148  
<211> 5  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> MOD\_RES

-125-

<222> 1  
<223> Xaa is Benzoyl-Val

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 148  
Xaa Gly Arg Ser Xaa  
1 5

<210> 149  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-HHT;  
HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 149  
Xaa Gly Arg Ser Xaa  
1 5

<210> 150  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-alpha-Z-D-Arg;  
Z is benzyloxycarbonyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 150  
Xaa Gly Arg Ser Xaa  
1 5

<210> 151  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate



-126-

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 151  
Xaa Gly Arg Ser Xaa  
1 5

<210> 152  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Ile

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 152  
Xaa Pro Arg Ser Xaa  
1 5

<210> 153  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Cbo-L-(gamma)Glu(alpha-t-BuO);  
Cbo is carbobenzoxy

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 153  
Xaa Gly Arg Ser Xaa  
1 5

<210> 154  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

-127-

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Pro

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 154  
Xaa Phe Arg Ser Xaa  
1 5

<210> 155  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 155  
Xaa Leu Arg Ser Xaa  
1 5

<210> 156  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile;  
Bz is benzoyl

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 156  
Xaa Glu Gly Arg Ser Xaa  
1 5

<210> 157  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

-128-

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile

<221> MOD\_RES  
<222> 2  
<223> Xaa is Glu(gamma-OMe)

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 157  
Xaa Xaa Gly Arg Ser Xaa  
1 5

<210> 158  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is benzoyle-Pro

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 158  
Xaa Phe Arg Ser Xaa  
1 5

<210> 159  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Phe

<221> MOD\_RES  
<222> 2  
<223> Xaa is Pip is pipercolinic acid

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic acid

<400> 159  
Xaa Xaa Arg Ser Xaa  
1 5

<210> 160

-129-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic acid

<400> 160  
Xaa Leu Lys Ser Xaa  
1 5

<210> 161  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Nle

<221> MOD\_RES  
<222> 2  
<223> Xaa is HHT: hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Xaa is leucine-therapeutic agent

<400> 161  
Xaa Xaa Lys Ser Xaa  
1 5

<210> 162  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 162  
Xaa Arg Thr Lys Arg Ser Xaa

```

1
5
<210> 163
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> MOD_RES
<222> 1
<223> Xaa is H-Arg

<221> MOD_RES
<222> 6
<223> Leucine-therapeutic agent

<400> 163
Xaa Gln Arg Arg Ser Xaa
 1 5

<210> 164
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> MOD_RES
<222> 1
<223> Xaa is Boc-Gln

<221> MOD_RES
<222> 5
<223> Xaa is Leucine-therapeutic agent

<400> 164
Xaa Gly Arg Ser Xaa
 1 5

<210> 165
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> MOD_RES
<222> 1
<223> Xaa is Z-Arg:
      Z is benzyloxycarbonyl

<221> MOD_RES
<222> 4
<223> Leucine-therapeutic agent

<400> 165
Xaa Arg Ser Xaa

```

-131-

1

<210> 166  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-HHT: HHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 166  
Xaa Ala Arg Ser Xaa  
1 5

<210> 167  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-CHT: CHT is hexahydrotyrosyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 167  
Xaa Gly Arg Ser Xaa  
1 5

<210> 168  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is MeSO<sub>2</sub>-dPhe

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 168  
Xaa Pro Arg Ser Xaa  
1 5

-132-

<210> 169  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is delta-Z-D-Lys: Z is benzyloxycarbonyl

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 169  
Xaa Pro Arg Ser Xaa  
1 5

<210> 170  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-CHA: CHA is cyclohexylalanyl

<221> MOD\_RES  
<222> 2  
<223> Xaa is But-Arg

<221> MOD\_RES  
<222> 4  
<223> Leucine-therapeutic agent

<400> 170  
Xaa Xaa Ser Xaa  
1

<210> 171  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 171

-133-

Arg Gln Ser Arg Ser Xaa  
1 5

<210> 172  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 172  
Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 173  
<211> 8  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 173  
Leu Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 174  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 174  
Arg Gln Ser Arg Xaa  
1 5

<210> 175



-134-

<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 175  
Arg Arg Gln Ser Arg Xaa  
1 5

<210> 176  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 176  
Leu Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 177  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 177  
Leu Arg Arg Gln Ser Arg Xaa  
1 5

<210> 178  
<211> 6  
<212> PRT  
<213> Artificial Sequence

-135-

<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 178  
Arg Arg Gln Ser Arg Xaa  
1 5  
  
<210> 179  
<211> 8  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent  
  
<400> 179  
Leu Arg Arg Gln Ser Arg Ser Xaa  
1 5  
  
<210> 180  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 180  
Arg Gln Gly Arg Ser Xaa  
1 5  
  
<210> 181  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION

-136-

<222> 1  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 181  
Arg Gln Ala Arg Ser Xaa  
1 5  
  
<210> 182  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 182  
Arg Gln Phe Arg Ser Xaa  
1 5  
  
<210> 183  
<211> 5  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent  
  
<400> 183  
Arg Ser Arg Ser Xaa  
1 5  
  
<210> 184  
<211> 5  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5

-137-

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 184

Arg Gly Arg Ser Xaa  
1 5

&lt;210&gt; 185

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 185

Arg Ala Arg Ser Xaa  
1 5

&lt;210&gt; 186

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 186

Arg Phe Arg Ser Xaa  
1 5

&lt;210&gt; 187

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 187

Gln Ser Arg Ser Xaa

-138-

1 5

<210> 188  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 188  
Gln Gly Arg Ser Xaa  
1 5

<210> 189  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 189  
Gln Ala Arg Ser Xaa  
1 5

<210> 190  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 190  
Gln Phe Arg Ser Xaa  
1 5

<210> 191  
<211> 9

-139-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(Alpha-(3-amidinobenzyl)

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 191  
Leu Arg Ala Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 192  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 192  
Leu Arg Ala Xaa Ala Arg Ser Ser Xaa  
1 5

<210> 193  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 9

-140-

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 193

Leu Arg Ser Xaa Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 194

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 194

Leu Arg Ser Xaa Ala Arg Ser Ser Xaa  
1 5

&lt;210&gt; 195

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 195

Leu Arg Pro Arg Phe Lys Ser Ser Xaa  
1 5

&lt;210&gt; 196

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

-141-

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 196

Arg Pro Arg Phe Lys Ser Ser Xaa  
1 5

&lt;210&gt; 197

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 197

Pro Arg Phe Lys Ser Ser Xaa  
1 5

&lt;210&gt; 198

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 198

Leu Arg Ser Lys Ser Arg Ser Ser Xaa  
1 5

&lt;210&gt; 199

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 199

Arg Ser Lys Ser Arg Ser Ser Xaa



-142-

1 5

<210> 200  
<211> 7  
<212> PRT  
<213> Aritificial sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 200  
Ser Lys Ser Arg Ser Ser Xaa  
1 5

<210> 201  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 201  
Leu Arg Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 202  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 202  
Arg Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 203  
<211> 7

-143-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 203  
Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 204  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 204  
Leu Arg Ser Arg Ser Arg Ser Ser Xaa  
1 5

<210> 205  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 205  
Arg Ser Arg Ser Arg Ser Ser Xaa  
1 5

<210> 206  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>

-144-

```

<223> Conjugate

<221> ACETYLATION
<222> 1

<221> MOD_RES
<222> 7
<223> Leucine-therapeutic agent

<400> 206
Ser Arg Ser Arg Ser Ser Xaa
 1             5

<210> 207
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> ACETYLATION
<222> (0)...(0)

<221> MOD_RES
<222> 4
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD_RES
<222> 9
<223> Leucine-therapeutic agent

<400> 207
Leu Arg Ala Xaa Gly Arg Ser Ser Xaa
 1             5

<210> 208
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> ACETYLATION
<222> 1

<221> MOD_RES
<222> 4
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD_RES
<222> 9
<223> Leucine-therapeutic agent

<400> 208
Leu Arg Ala Xaa Ala Arg Ser Ser Xaa
 1             5

<210> 209
<211> 9

```

-145-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 209  
Leu Arg Ser Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 210  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is Quat: (R)-Glu(alpha-(3-amidinobenzyl))

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 210  
Leu Arg Ser Xaa Ala Arg Ser Ser Xaa  
1 5

<210> 211  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 211  
Leu Arg Pro Arg Phe Lys Ser Ser Xaa

-146-

1 5

<210> 212  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 212  
Arg Pro Arg Phe Lys Ser Ser Xaa  
1 5

<210> 213  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 213  
Pro Arg Phe Lys Ser Ser Xaa  
1 5

<210> 214  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 214  
Leu Arg Ser Lys Ser Arg Ser Ser Xaa  
1 5

<210> 215  
<211> 8

-147-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 215  
Arg Ser Lys Ser Arg Ser Ser Xaa  
1 5

<210> 216  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 216  
Ser Lys Ser Arg Ser Ser Xaa  
1 5

<210> 217  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 217  
Leu Arg Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 218  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>

-148-

<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 218  
Arg Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 219  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 219  
Pro Arg Phe Arg Ser Ser Xaa  
1 5

<210> 220  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 220  
Leu Arg Ser Arg Ser Arg Ser Ser Xaa  
1 5

<210> 221  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

-149-

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 221  
Arg Ser Arg Ser Arg Ser Ser Xaa  
1 5

<210> 222  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 222  
Ser Arg Ser Arg Ser Ser Xaa  
1 5

<210> 223  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 223  
Xaa Pro Arg Ser Ser Xaa  
1 5

<210> 224  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-HHT;  
HHT is hexahydrotyrosyl

<221> MOD\_RES



-150-

<222> 6  
<223> Leucine-therapeutic agent

<400> 224  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 225  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is n-p-tosyl-Gly

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 225  
Xaa Pro Arg Ser Ser Xaa  
1 5

<210> 226  
<211> 6  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Benzoyl-Val

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 226  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 227  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-HHT;  
HHT is hexahydrotyrosyl

<221> MOD\_RES

-151-

<222> 6  
<223> Leucine-therapeutic agent

<400> 227  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 228  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-alpha-Z-D-Arg;  
Z is benzyloxycarbonyl

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 228  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 229  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 229  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 230  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Ile

<221> MOD\_RES

-152-

<222> 6  
<223> Leucine-therapeutic agent

<400> 230  
Xaa Pro Arg Ser Ser Xaa  
1 5

<210> 231  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Cbo-L-(gamma)Glu(alpha-t-BuO);  
Cbo is carbobenzoxy

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 231  
Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 232  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Pro

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 232  
Xaa Phe Arg Ser Ser Xaa  
1 5

<210> 233  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MOD\_RES

-153-

<222> 6  
<223> Leucine-therapeutic agent

<400> 233  
Xaa Leu Arg Ser Ser Xaa  
1 5

<210> 234  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile;  
Bz is benzoyl

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 234  
Xaa Glu Gly Arg Ser Ser Xaa  
1 5

<210> 235  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Bz-Ile

<221> MOD\_RES  
<222> 2  
<223> Xaa is Glu(gamma-OMe)

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 235  
Xaa Xaa Gly Arg Ser Ser Xaa  
1 5

<210> 236  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES

-154-

<222> 1  
<223> Xaa is benzoyle-Pro

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 236  
Xaa Phe Arg Ser Ser Xaa  
1 5

<210> 237  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Phe

<221> MOD\_RES  
<222> 2  
<223> Xaa is Pip is pipecolinic acid

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic acid

<400> 237  
Xaa Xaa Arg Ser Ser Xaa  
1 5

<210> 238  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Val

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic acid

<400> 238  
Xaa Leu Lys Ser Ser Xaa  
1 5

<210> 239  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>

-155-

<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-D-Nle

<221> MOD\_RES  
<222> 2  
<223> Xaa is HHT: hexahydrotyrosyl

<221> MOD\_RES  
<222> 6  
<223> Xaa is leucine-therapeutic agent

<400> 239  
Xaa Xaa Lys Ser Ser Xaa  
1 5

<210> 240  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is pyroglutamic acid

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 240  
Xaa Arg Thr Lys Arg Ser Ser Xaa  
1 5

<210> 241  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is H-Arg

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 241  
Xaa Gln Arg Arg Ser Ser Xaa  
1 5

<210> 242  
<211> 6  
<212> PRT

-156-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Boc-Gln

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Xaa is Leucine-therapeutic agent

&lt;400&gt; 242

Xaa Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 243

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is Z-Arg:  
Z is benzyloxycarbonyl

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 243

Xaa Arg Ser Ser Xaa  
1 5

&lt;210&gt; 244

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is H-D-HHT: HHT is hexahydrotyrosyl

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 244

Xaa Ala Arg Ser Ser Xaa  
1 5

&lt;210&gt; 245

&lt;211&gt; 6

&lt;212&gt; PRT

-157-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is H-D-CHT: CHT is hexahydrotyrosyl

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 245

Xaa Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 246

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is MeSO<sub>2</sub>-dPhe

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 246

Xaa Pro Arg Ser Ser Xaa  
1 5

&lt;210&gt; 247

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is delta-Z-D-Lys: Z is benzyloxycarbonyl

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 247

Xaa Pro Arg Ser Ser Xaa  
1 5

&lt;210&gt; 248

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence



-158-

<220>  
<223> Conjugate  
  
<221> MOD\_RES  
<222> 1  
<223> Xaa is CH<sub>3</sub>SO<sub>2</sub>-D-CHA: CHA is cyclohexylalanyl

<221> MOD\_RES  
<222> 2  
<223> Xaa is But-Arg

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 248  
Xaa Xaa Ser Ser Xaa  
1 5

<210> 249  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 249  
Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 250  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 250  
Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 251  
<211> 9  
<212> PRT

-159-

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 251

Leu Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

&lt;210&gt; 252

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 252

Arg Gln Ser Arg Xaa  
1 5

&lt;210&gt; 253

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 253

Arg Arg Gln Ser Arg Xaa  
1 5

&lt;210&gt; 254

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Conjugate

-160-

<221> MOD\_RES  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic agent

<400> 254  
Leu Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 255  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 255  
Leu Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 256  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 256  
Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 257  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES

-161-

<222> 9  
<223> Leucine-therapeutic agent

<400> 257  
Leu Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 258  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 258  
Arg Gln Gly Arg Ser Ser Xaa  
1 5

<210> 259  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 259  
Arg Gln Ala Arg Ser Ser Xaa  
1 5

<210> 260  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic acid

<400> 260

-162-

Arg Gln Phe Arg Ser Ser Xaa  
1 5

<210> 261  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 261  
Arg Ser Arg Ser Ser Xaa  
1 5

<210> 262  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 262  
Arg Gly Arg Ser Ser Xaa  
1 5

<210> 263  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 263  
Arg Ala Arg Ser Ser Xaa  
1 5

<210> 264

-163-

<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 264  
Arg Phe Arg Ser Ser Xaa  
1 5

<210> 265  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 265  
Gln Ser Arg Ser Ser Xaa  
1 5

<210> 266  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 266  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 267  
<211> 6  
<212> PRT  
<213> Artificial Sequence

-164-

```

<220>
<223> Conjugate

<221> ACETYLATION
<222> 1

<221> MOD_RES
<222> 6
<223> leucine-therapeutic agent

<400> 267
Gln Ala Arg Ser Ser Xaa
 1             5

<210> 268
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Conjugate

<221> ACETYLATION
<222> 1

<221> MOD_RES
<222> 6
<223> Leucine-therapeutic agent

<400> 268
Gln Phe Arg Ser Ser Xaa
 1             5

<210> 269
<211> 816
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (1)...(816)
<223> Nucleotide sequence encoding MTSP25, including
      MTSP25 protease domain

<221> misc feature
<222> (248)...(270)
<223> Transmembrane domain encompasses amino acids
      248-270 at the C-terminus of the trypsin-like
      serine protease domain (amino acids 1-237)

<400> 269
att ata ggg ggc acc gaa gca caa gct ggc gca tgg ccg tgg gtg gtg   48
Ile Ile Gly Gly Thr Glu Ala Gln Ala Gly Ala Trp Pro Trp Val Val
 1             5             10             15

agc ctg cag att aaa tat ggc cgt gtt ctt gtt cat gta tgt ggg gga   96
Ser Leu Gln Ile Lys Tyr Gly Arg Val Leu Val His Val Cys Gly Gly
          20             25             30

acc cta gtg aga gag agg tgg gtc ctc aca gct gcc cac tgc act aaa   144
Thr Leu Val Arg Glu Arg Trp Val Leu Thr Ala Ala His Cys Thr Lys

```

35				40				45								
gac Asp	gct Ala 50	agc Ser	gat Asp	cct Pro	tta Leu	atg Met 55	tgg Trp	aca Thr	gct Ala	gtg Val	att Ile 60	gga Gly	act Thr	aat Asn	aat Asn	192
ata Ile 65	cat His	gga Gly	cgc Arg	tat Tyr	cct Pro 70	cat His	acc Thr	aag Lys	aag Lys	ata Ile 75	aaa Lys	att Ile	aaa Lys	gca Ala	atc Ile 80	240
att Ile	att Ile	cat His	cca Pro	aac Asn 85	ttc Phe	att Ile	ttg Leu	gaa Glu	tct Ser 90	tat Tyr	gta Val	aat Asn	gat Asp	att Ile 95	gca Ala	288
ctt Leu	ttt Phe	cac His	tta Leu 100	aaa Lys	aaa Lys	gca Ala	gtg Val	agg Arg 105	tat Tyr	aat Asn	gac Asp	tat Tyr	att Ile 110	cag Gln	cct Pro	336
att Ile	tgc Cys 115	cta Leu	cct Pro	ttt Phe	gat Asp	gtt Val	ttc Phe 120	caa Gln	atc Ile	ctg Leu	gac Asp	gga Gly 125	aac Asn	aca Thr	aag Lys	384
tgt Cys 130	ttt Phe	ata Ile	agt Ser	ggc Gly	tgg Trp	gga Gly 135	aga Arg	aca Thr	aaa Lys	gaa Glu	gaa Glu 140	ggg Gly	aac Asn	gct Ala	aca Thr	432
aat Asn 145	att Ile	tta Leu	caa Gln	gat Asp	gca Ala 150	gaa Glu	gtg Val	cat His	tat Tyr	att Ile 155	tct Ser	cga Arg	gag Glu	atg Met	tgt Cys 160	480
aat Asn	tct Ser	gag Glu	agg Arg	agt Ser 165	tat Tyr	ggg Gly	gga Gly	ata Ile	att Ile 170	cct Pro	aac Asn	act Thr	tca Ser	ttt Phe 175	tgt Cys	528
gca Ala	ggg Gly	gat Asp	gaa Glu 180	gat Asp	gga Gly	gct Ala	ttt Phe 185	gat Asp	act Thr	tgc Cys	agg Arg	ggg Gly	gac Asp 190	agt Ser	ggg Gly	576
gga Gly	cca Pro	tta Leu 195	atg Met	tgc Cys	tac Tyr	tta Leu 200	cca Pro	gaa Glu	tat Tyr	aaa Lys	aga Arg	ttt Phe 205	ttt Phe	gta Val	atg Met	624
gga Gly	att Ile 210	acc Thr	agt Ser	tac Tyr	gga Gly	cat His 215	ggc Gly	tgt Cys	ggg Gly	cga Arg	aga Arg 220	ggg Gly	ttt Phe	cct Pro	ggg Gly	672
gtc Val 225	tat Tyr	att Ile	ggg Gly	cca Pro	tcc Ser 230	ttc Phe	tac Tyr	caa Gln	aag Lys	tgg Trp 235	ctg Leu	aca Thr	gag Glu	cat His	ttc Phe 240	720
ttc Phe	cat His	gca Ala	agc Ser	act Thr 245	caa Gln	ggc Gly	ata Ile	ctt Leu	act Thr 250	ata Ile	aat Asn	att Ile	tta Leu	cgt Arg 255	ggc Gly	768
cag Gln	atc Ile	ctc Leu	ata Ile 260	gct Ala	tta Leu	tgt Cys	ttt Phe 265	gtc Val	atc Ile	tta Leu	cta Leu	gca Ala	aca Thr 270	aca Thr	taa *	816
<210> 270																
<211> 271																
<212> PRT																



-166-

&lt;213&gt; Homo Sapien

&lt;400&gt; 270

```

Ile Ile Gly Gly Thr Glu Ala Gln Ala Gly Ala Trp Pro Trp Val Val
 1      5      10      15
Ser Leu Gln Ile Lys Tyr Gly Arg Val Leu Val His Val Cys Gly Gly
      20      25      30
Thr Leu Val Arg Glu Arg Trp Val Leu Thr Ala Ala His Cys Thr Lys
      35      40      45
Asp Ala Ser Asp Pro Leu Met Trp Thr Ala Val Ile Gly Thr Asn Asn
 50      55      60
Ile His Gly Arg Tyr Pro His Thr Lys Lys Ile Lys Ile Lys Ala Ile
65      70      75      80
Ile Ile His Pro Asn Phe Ile Leu Glu Ser Tyr Val Asn Asp Ile Ala
      85      90      95
Leu Phe His Leu Lys Lys Ala Val Arg Tyr Asn Asp Tyr Ile Gln Pro
      100      105      110
Ile Cys Leu Pro Phe Asp Val Phe Gln Ile Leu Asp Gly Asn Thr Lys
      115      120      125
Cys Phe Ile Ser Gly Trp Gly Arg Thr Lys Glu Glu Gly Asn Ala Thr
      130      135      140
Asn Ile Leu Gln Asp Ala Glu Val His Tyr Ile Ser Arg Glu Met Cys
145      150      155      160
Asn Ser Glu Arg Ser Tyr Gly Gly Ile Ile Pro Asn Thr Ser Phe Cys
      165      170      175
Ala Gly Asp Glu Asp Gly Ala Phe Asp Thr Cys Arg Gly Asp Ser Gly
      180      185      190
Gly Pro Leu Met Cys Tyr Leu Pro Glu Tyr Lys Arg Phe Phe Val Met
      195      200      205
Gly Ile Thr Ser Tyr Gly His Gly Cys Gly Arg Arg Gly Phe Pro Gly
      210      215      220
Val Tyr Ile Gly Pro Ser Phe Tyr Gln Lys Trp Leu Thr Glu His Phe
225      230      235      240
Phe His Ala Ser Thr Gln Gly Ile Leu Thr Ile Asn Ile Leu Arg Gly
      245      250      255
Gln Ile Leu Ile Ala Leu Cys Phe Val Ile Leu Leu Ala Thr Thr
      260      265      270

```

&lt;210&gt; 271

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; amino acids 401-407 of SEQ ID No. 97 in WO 02/00860

&lt;400&gt; 271

```

Arg Lys His Leu Pro Arg Pro Ala
 1      5

```

&lt;210&gt; 272

&lt;211&gt; 228

&lt;212&gt; PRT

&lt;213&gt; alternative PD1 of MTSP12

&lt;400&gt; 272

```

Ile Val Gly Gly Met Glu Ala Ser Pro Gly Glu Phe Pro Trp Gln Ala
 1      5      10      15
Ser Leu Arg Glu Asn Lys Glu His Phe Cys Gly Ala Ala Ile Ile Asn
      20      25      30
Ala Arg Trp Leu Val Ser Ala Ala His Cys Phe Asn Glu Phe Gln Asp
      35      40      45

```

-167-

```

Pro Thr Lys Trp Val Ala Tyr Val Gly Ala Thr Tyr Leu Ser Gly Ser
  50      55      60
Glu Ala Ser Thr Val Arg Ala Gln Val Val Gln Ile Val Lys His Pro
  65      70      75      80
Leu Tyr Asn Ala Asp Thr Ala Asp Phe Asp Val Ala Val Leu Glu Leu
      85      90      95
Thr Ser Pro Leu Pro Phe Gly Arg His Ile Gln Pro Val Cys Leu Pro
      100      105      110
Ala Ala Thr His Ile Phe Pro Pro Ser Lys Lys Cys Leu Ile Ser Gly
      115      120      125
Trp Gly Tyr Leu Lys Glu Asp Phe Leu Arg Lys His Leu Pro Arg Pro
      130      135      140
Ala Val Lys Pro Gly Val Leu Gln Lys Ala Thr Val Glu Leu Leu Asp
  145      150      155      160
Gln Ala Leu Cys Ala Ser Leu Tyr Gly His Ser Leu Thr Asp Arg Met
      165      170      175
Val Cys Ala Gly Tyr Leu Asp Gly Lys Val Asp Ser Cys Gln Gly Asp
      180      185      190
Ser Gly Gly Pro Leu Val Cys Glu Glu Pro Ser Gly Arg Phe Ser Leu
      195      200      205
Ala Gly Ile Val Ser Trp Gly Ile Gly Cys Ala Glu Ala Arg Arg Pro
      210      215      220
Gly Val Tyr Ala
  225

```

&lt;210&gt; 273

&lt;211&gt; 804

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(804)

<223> Nucleotide sequence encoding MTSP20, including  
MTSP20 protease domain

&lt;400&gt; 273

```

aca gca ggt ccc cag gca gga gca ccc tcc cca tgg ccc tgg gag gcc      48
Thr Ala Gly Pro Gln Ala Gly Ala Pro Ser Pro Trp Pro Trp Glu Ala
  1      5      10      15

agg ctg atg cac cag gga cag ctg gcc tgt ggc gga gcc ctg gtg tca      96
Arg Leu Met His Gln Gly Gln Leu Ala Cys Gly Gly Ala Leu Val Ser
      20      25      30

gag gag acg gtg cta act gtt gcc cac tgc ttc att ggg cgc cag gcc      144
Glu Glu Thr Val Leu Thr Val Ala His Cys Phe Ile Gly Arg Gln Ala
      35      40      45

cca gag gaa tgg agc gta ggg ctg ggg acc aga ccg gag gag tgg ggc      192
Pro Glu Glu Trp Ser Val Gly Leu Gly Thr Arg Pro Glu Glu Trp Gly
      50      55      60

ctg aag cag ctc atc ctg cat gga gcc tac acc cac cct gag ggg ggc      240
Leu Lys Gln Leu Ile Leu His Gly Ala Tyr Thr His Pro Glu Gly Gly
      65      70      75      80

tac gac atg gcc ctc ctg ctg ctg gcc cag cct gtg aca ctg gga gcc      288
Tyr Asp Met Ala Leu Leu Leu Leu Ala Gln Pro Val Thr Leu Gly Ala
      85      90      95

```

-168-

agc	ctg	cgg	ccc	ctc	tgc	ctg	ccc	tat	cct	gac	cac	cac	ctg	cct	gat	336
Ser	Leu	Arg	Pro	Leu	Cys	Leu	Pro	Tyr	Pro	Asp	His	His	Leu	Pro	Asp	
			100					105					110			
ggg	gag	cgt	ggc	tgg	gtt	ctg	gga	cgg	gcc	cgc	cca	gga	gca	ggc	atc	384
Gly	Glu	Arg	Gly	Trp	Val	Leu	Gly	Arg	Ala	Arg	Pro	Gly	Ala	Gly	Ile	
		115					120					125				
agc	tcc	ctc	cag	aca	gtg	ccc	gtg	acc	ctc	ctg	ggg	cct	agg	gcc	tgc	432
Ser	Ser	Leu	Gln	Thr	Val	Pro	Val	Thr	Leu	Leu	Gly	Pro	Arg	Ala	Cys	
		130				135					140					
agc	cgg	ctg	cat	gca	gct	cct	ggg	ggt	gat	ggc	agc	cct	att	ctg	ccg	480
Ser	Arg	Leu	His	Ala	Ala	Pro	Gly	Gly	Asp	Gly	Ser	Pro	Ile	Leu	Pro	
		145			150					155					160	
ggg	atg	gtg	tgt	acc	agt	gct	gtg	ggt	gag	ctg	ccc	agc	tgt	gag	ggc	528
Gly	Met	Val	Cys	Thr	Ser	Ala	Val	Gly	Glu	Leu	Pro	Ser	Cys	Glu	Gly	
				165					170					175		
ctg	tct	ggg	gca	cca	ctg	gtg	cat	gag	gtg	agg	ggc	aca	tgg	ttc	ctg	576
Leu	Ser	Gly	Ala	Pro	Leu	Val	His	Glu	Val	Arg	Gly	Thr	Trp	Phe	Leu	
			180					185					190			
gcc	ggg	ctg	cac	agc	ttc	gga	gat	gct	tgc	caa	ggc	ccc	gcc	agg	ccg	624
Ala	Gly	Leu	His	Ser	Phe	Gly	Asp	Ala	Cys	Gln	Gly	Pro	Ala	Arg	Pro	
		195					200					205				
gcg	gtc	ttc	acc	gcg	ctc	cct	gcc	tat	gag	gac	tgg	gtc	agc	agt	ttg	672
Ala	Val	Phe	Thr	Ala	Leu	Pro	Ala	Tyr	Glu	Asp	Trp	Val	Ser	Ser	Leu	
		210				215					220					
gac	tgg	cag	gtc	tac	ttc	gcc	gag	gaa	cca	gag	ccc	gag	gct	gag	cct	720
Asp	Trp	Gln	Val	Tyr	Phe	Ala	Glu	Glu	Pro	Glu	Pro	Glu	Ala	Glu	Pro	
		225			230					235					240	
gga	agc	tgc	ctg	gcc	aac	atg	agt	atg	tgg	ccc	cgg	ggc	ctc	ctg	cca	768
Gly	Ser	Cys	Leu	Ala	Asn	Met	Ser	Met	Trp	Pro	Arg	Gly	Leu	Leu	Pro	
				245					250					255		
aac	cct	gcc	tct	cca	gga	ccc	ttc	tct	ctc	cag	tga					804
Asn	Pro	Ala	Ser	Pro	Gly	Pro	Phe	Ser	Leu	Gln	*					
			260				265									

&lt;210&gt; 274

&lt;211&gt; 267

&lt;212&gt; PRT

&lt;213&gt; Homo Sapien

&lt;400&gt; 274

Thr	Ala	Gly	Pro	Gln	Ala	Gly	Ala	Pro	Ser	Pro	Trp	Pro	Trp	Glu	Ala
1				5					10					15	
Arg	Leu	Met	His	Gln	Gly	Gln	Leu	Ala	Cys	Gly	Gly	Ala	Leu	Val	Ser
			20					25					30		
Glu	Glu	Thr	Val	Leu	Thr	Val	Ala	His	Cys	Phe	Ile	Gly	Arg	Gln	Ala
		35				40						45			
Pro	Glu	Glu	Trp	Ser	Val	Gly	Leu	Gly	Thr	Arg	Pro	Glu	Glu	Trp	Gly
		50				55					60				
Leu	Lys	Gln	Leu	Ile	Leu	His	Gly	Ala	Tyr	Thr	His	Pro	Glu	Gly	Gly
65					70				75					80	

-169-

Tyr Asp Met Ala Leu Leu Leu Leu Ala Gln Pro Val Thr Leu Gly Ala  
 85 90 95  
 Ser Leu Arg Pro Leu Cys Leu Pro Tyr Pro Asp His His Leu Pro Asp  
 100 105 110  
 Gly Glu Arg Gly Trp Val Leu Gly Arg Ala Arg Pro Gly Ala Gly Ile  
 115 120 125  
 Ser Ser Leu Gln Thr Val Pro Val Thr Leu Leu Gly Pro Arg Ala Cys  
 130 135 140  
 Ser Arg Leu His Ala Ala Pro Gly Gly Asp Gly Ser Pro Ile Leu Pro  
 145 150 155 160  
 Gly Met Val Cys Thr Ser Ala Val Gly Glu Leu Pro Ser Cys Glu Gly  
 165 170 175  
 Leu Ser Gly Ala Pro Leu Val His Glu Val Arg Gly Thr Trp Phe Leu  
 180 185 190  
 Ala Gly Leu His Ser Phe Gly Asp Ala Cys Gln Gly Pro Ala Arg Pro  
 195 200 205  
 Ala Val Phe Thr Ala Leu Pro Ala Tyr Glu Asp Trp Val Ser Ser Leu  
 210 215 220  
 Asp Trp Gln Val Tyr Phe Ala Glu Glu Pro Glu Pro Glu Ala Glu Pro  
 225 230 235 240  
 Gly Ser Cys Leu Ala Asn Met Ser Met Trp Pro Arg Gly Leu Leu Pro  
 245 250 255  
 Asn Pro Ala Ser Pro Gly Pro Phe Ser Leu Gln  
 260 265

&lt;210&gt; 275

&lt;211&gt; 699

&lt;212&gt; DNA

&lt;213&gt; Homo Sapien

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(699)

 <223> Nucleotide sequence encoding MTSP22, including  
 MTSP22 protease domain

&lt;400&gt; 275

att gtg aat gga aaa agc tcc ctg gag ggg gca tgg cca tgg cag gcc	48
Ile Val Asn Gly Lys Ser Ser Leu Glu Gly Ala Trp Pro Trp Gln Ala	
1 5 10 15	
agc atg caa tgg aaa ggc cgt cac tac tgt gga gcc tct ctg atc agc	96
Ser Met Gln Trp Lys Gly Arg His Tyr Cys Gly Ala Ser Leu Ile Ser	
20 25 30	
agc agg tgg cta tta tct gca gct cac tgc ttt gct aag aaa aat aat	144
Ser Arg Trp Leu Leu Ser Ala Ala His Cys Phe Ala Lys Lys Asn Asn	
35 40 45	
tca aaa gat tgg act gtc aac ttt gga gtt gta gta aat aaa cca tat	192
Ser Lys Asp Trp Thr Val Asn Phe Gly Val Val Val Asn Lys Pro Tyr	
50 55 60	
atg aca cgg aaa gtc caa aac att att ttt cat gaa aat tat agc agt	240
Met Thr Arg Lys Val Gln Asn Ile Ile Phe His Glu Asn Tyr Ser Ser	
65 70 75 80	
cct ggg ctt cat gat gat att gcc ctt gtg cag ctt gct gaa gaa gtt	288
Pro Gly Leu His Asp Asp Ile Ala Leu Val Gln Leu Ala Glu Glu Val	
85 90 95	

-170-

tct ttt aca gag tac att cgt aag att tgt ctt cct gaa gcc aaa atg	336
Ser Phe Thr Glu Tyr Ile Arg Lys Ile Cys Leu Pro Glu Ala Lys Met	
100 105 110	
aag ctc tca gaa aat gac aat gtt gta gtt aca ggt tgg gga aca ctt	384
Lys Leu Ser Glu Asn Asp Asn Val Val Val Thr Gly Trp Gly Thr Leu	
115 120 125	
tat atg aat ggt tca ttt cca gtg ata ctt caa gaa gcc ttt ttg aag	432
Tyr Met Asn Gly Ser Phe Pro Val Ile Leu Gln Glu Ala Phe Leu Lys	
130 135 140	
att att gac aac aaa att tgc aat gcc tca tat gca tac tct ggc tta	480
Ile Ile Asp Asn Lys Ile Cys Asn Ala Ser Tyr Ala Tyr Ser Gly Leu	
145 150 155 160	
gtg act gat aca atg tta tgt gct gga ttt atg tca gga gaa gct gat	528
Val Thr Asp Thr Met Leu Cys Ala Gly Phe Met Ser Gly Glu Ala Asp	
165 170 175	
gca tgt cag aat gat tct ggt gga cca cta gct tac cct gat tcc aga	576
Ala Cys Gln Asn Asp Ser Gly Gly Pro Leu Ala Tyr Pro Asp Ser Arg	
180 185 190	
aat atc tgg cat ctt gtt gga ata gta agc tgg ggt gat gga tgt ggt	624
Asn Ile Trp His Leu Val Gly Ile Val Ser Trp Gly Asp Gly Cys Gly	
195 200 205	
aaa aag aat aag cca ggt gtc tat act cga gtg act tct tat cgc aat	672
Lys Lys Asn Lys Pro Gly Val Tyr Thr Arg Val Thr Ser Tyr Arg Asn	
210 215 220	
tgg att aca tcc aag act gga ctc tga	699
Trp Ile Thr Ser Lys Thr Gly Leu *	
225 230	

<210> 276  
 <211> 232  
 <212> PRT  
 <213> Homo Sapien

<400> 276  
 Ile Val Asn Gly Lys Ser Ser Leu Glu Gly Ala Trp Pro Trp Gln Ala  
 1 5 10 15  
 Ser Met Gln Trp Lys Gly Arg His Tyr Cys Gly Ala Ser Leu Ile Ser  
 20 25 30  
 Ser Arg Trp Leu Leu Ser Ala Ala His Cys Phe Ala Lys Lys Asn Asn  
 35 40 45  
 Ser Lys Asp Trp Thr Val Asn Phe Gly Val Val Val Asn Lys Pro Tyr  
 50 55 60  
 Met Thr Arg Lys Val Gln Asn Ile Ile Phe His Glu Asn Tyr Ser Ser  
 65 70 75 80  
 Pro Gly Leu His Asp Asp Ile Ala Leu Val Gln Leu Ala Glu Glu Val  
 85 90 95  
 Ser Phe Thr Glu Tyr Ile Arg Lys Ile Cys Leu Pro Glu Ala Lys Met  
 100 105 110  
 Lys Leu Ser Glu Asn Asp Asn Val Val Val Thr Gly Trp Gly Thr Leu  
 115 120 125  
 Tyr Met Asn Gly Ser Phe Pro Val Ile Leu Gln Glu Ala Phe Leu Lys  
 130 135 140

-171-

Ile	Ile	Asp	Asn	Lys	Ile	Cys	Asn	Ala	Ser	Tyr	Ala	Tyr	Ser	Gly	Leu
145					150					155					160
Val	Thr	Asp	Thr	Met	Leu	Cys	Ala	Gly	Phe	Met	Ser	Gly	Glu	Ala	Asp
				165					170						175
Ala	Cys	Gln	Asn	Asp	Ser	Gly	Gly	Pro	Leu	Ala	Tyr	Pro	Asp	Ser	Arg
			180					185					190		
Asn	Ile	Trp	His	Leu	Val	Gly	Ile	Val	Ser	Trp	Gly	Asp	Gly	Cys	Gly
		195					200					205			
Lys	Lys	Asn	Lys	Pro	Gly	Val	Tyr	Thr	Arg	Val	Thr	Ser	Tyr	Arg	Asn
	210					215					220				
Trp	Ile	Thr	Ser	Lys	Thr	Gly	Leu								

<210> 277  
 <211> 6  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> ACETYLATION  
 <222> 1

<221> MOD\_RES  
 <222> 6  
 <223> Leucine-therapeutic agent  
 <223> conjugate

<400> 277  
 Gly Ser Gly Arg Ser Xaa  
 1 5

<210> 278  
 <211> 6  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> ACETYLATION  
 <222> 1

<221> MOD\_RES  
 <222> 6  
 <223> Xaa is Leucine-therapeutic Agent  
 <223> conjugate

<400> 278  
 Gly Ser Gly Arg Ser Xaa  
 1 5

<210> 279  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <221> ACETYLATION  
 <222> 1

-172-

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent  
  
<223> conjugate

<400> 279  
Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 280  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic Agent  
  
<223> conjugate

<400> 280  
Gly Ser Gly Arg Xaa  
1 5

<210> 281  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> AMIDATION  
<222> 6

<221> MOD\_RES  
<222> 4  
<223> Xaa is 4-Guanidino-phenylglycine

<223> conjugate

<400> 281  
Gly Ser Gly Xaa Ser Leu  
1 5

<210> 282  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION

-173-

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Cyclohexylamine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 282

Gly Ser Gly Arg Ser Ser Xaa

1

5

&lt;210&gt; 283

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 283

Gly Ser Gly Arg Ala Ser Xaa

1

5

&lt;210&gt; 284

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 284

Gly Ser Gly Arg Ser Xaa

1

5

&lt;210&gt; 285

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; (0)...(0)



-174-

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<223> conjugate

<400> 285  
Gly Thr Gly Arg Ser Xaa  
1 5

<210> 286  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Succinyl-BAlanine

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic Agent

<223> conjugate

<400> 286  
Ala Thr Gly Arg Ser Xaa  
1 5

<210> 287  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 287  
Gly Thr Gly Arg Ser Xaa  
1 5

<210> 288  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

-175-

<221> MOD\_RES  
<222> 2  
<223> Xaa is Homoserine

<221> MOD\_RES  
<222> 6  
<223> Nle-Therapeutic AgentNle

<223> conjugate

<400> 288  
Gly Xaa Gly Arg Ser Xaa  
1 5

<210> 289  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is D Serine

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 289  
Gly Xaa Ala Arg Ser Xaa  
1 5

<210> 290  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 290  
Gly Ser Ala Arg Ser Xaa  
1 5

<210> 291  
<211> 7

-176-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 291  
Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 292  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 292  
Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 293  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 293  
Gly Ser Ala Arg Ala Ser Xaa  
1 5

<210> 294  
<211> 6  
<212> PRT

-177-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 294

Val Ser Gly Arg Ser Xaa  
1 5

&lt;210&gt; 295

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 295

Val Ser Gly Arg Ala Xaa  
1 5

&lt;210&gt; 296

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 296

Val Ser Gly Arg Ala Ser Xaa  
1 5

&lt;210&gt; 297

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-178-

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 297  
Val Ser Gly Arg Ser Ser Xaa  
1 5

<210> 298  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 298  
Val Ser Ala Arg Met Xaa  
1 5

<210> 299  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Xaa is Nle

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 299  
Val Ser Ala Arg Xaa Xaa  
1 5

<210> 300  
<211> 6

-179-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<223> conjugate

<400> 300  
Val Ser Ala Arg Ser Xaa  
1 5

<210> 301  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 301  
Val Ser Ala Arg Ser Xaa  
1 5

<210> 302  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 6  
<223> dValine-therapeutic Agent

<223> conjugate

<400> 302  
Xaa Pro Gly Arg Val Xaa  
1 5

-180-

<210> 303  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 6  
<223> Valine-therapeutic Agent

<223> conjugate

<400> 303  
Xaa Pro Gly Arg Val Xaa  
1 5

<210> 304  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 304  
Xaa Pro Gly Arg Ala Xaa  
1 5

<210> 305  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

-181-

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 305  
Xaa Pro Gly Arg Ser Xaa  
1 5

<210> 306  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 306  
Xaa Pro Ala Arg Ser Xaa  
1 5

<210> 307  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 307  
Xaa Pro Ala Arg Ala Ser Xaa  
1 5

<210> 308  
<211> 6



-182-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is t-Butyl Glycine

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 308  
Xaa Pro Gly Arg Ser Xaa  
1 5

<210> 309  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is D Serine

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 309  
Arg Gly Xaa Ala Arg Ser Xaa  
1 5

<210> 310  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic Agent

<223> conjugate

-183-

<400> 310  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 311  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 311  
Arg Gly Ser Gly Arg Ala Xaa  
1 5

<210> 312  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 312  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 313  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nle-therapeutic agent

<223> conjugate

<400> 313

-184-

Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 314  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nle-therapeutic agent

<223> conjugate

<400> 314  
Arg Gly Ser Gly Arg Ala Xaa  
1 5

<210> 315  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 315  
Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 316  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 316  
Arg Gly Ser Gly Arg Ser Xaa

-185-

1

5

<210> 317  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Cyclohexylalanine-therapeutic agent  
  
<223> conjugate

<400> 317  
Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 318  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Cyclohexylalanine-therapeutic agent  
  
<223> conjugate

<400> 318  
Arg Gly Ser Ala Arg Ser Xaa  
1 5

<210> 319  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Cyclohexylalanine-therapeutic agent  
  
<223> conjugate

<400> 319  
Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

-186-

<210> 320  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nva-therapeutic agent  
  
<223> conjugate

<400> 320  
Arg Gly Ser Ala Arg Ser Xaa  
1 5

<210> 321  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Nva-therapeutic agent  
  
<223> conjugate

<400> 321  
Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 322  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent  
  
<223> conjugate

<400> 322  
Arg Gly Ser Ala Arg Ser Xaa  
1 5

-187-

<210> 323  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is S-MethylCysteine

<221> MOD\_RES  
<222> 7  
<223> Valine-therapeutic Agent

<223> conjugate

<400> 323  
Arg Xaa Pro Gly Arg Val Xaa  
1 5

<210> 324  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 7  
<223> Valine-therapeutic Agent

<223> conjugate

<400> 324  
Arg Xaa Pro Gly Arg Val Xaa  
1 5

<210> 325  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is S-Methylcysteine

-188-

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 325  
Arg Xaa Pro Gly Arg Ser Xaa  
1 5

<210> 326  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 326  
Arg Leu Pro Gly Arg Ser Xaa  
1 5

<210> 327  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> VARIANT  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 327  
Arg Val Pro Gly Arg Ser Xaa  
1 5

<210> 328  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> VARIANT

-189-

<222> 8  
<223> dLeucine-therapeutic Agent

<223> conjugate

<400> 328  
Arg Val Pro Gly Arg Ser Xaa  
1 5

<210> 329  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Nle

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 329  
Arg Xaa Pro Gly Arg Ser Xaa  
1 5

<210> 330  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is t-Butylglycine

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 330  
Arg Xaa Pro Ala Arg Ser Xaa  
1 5

<210> 331  
<211> 7  
<212> PRT



-190-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 331

Arg Leu Pro Ala Arg Ser Xaa  
1 5

&lt;210&gt; 332

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 332

Arg Val Pro Ala Arg Ser Xaa  
1 5

&lt;210&gt; 333

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is Nle

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 333

Arg Xaa Pro Ala Arg Ser Xaa  
1 5

-191-

<210> 334  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 334  
Ile Val Ser Gly Arg Ala Xaa  
1 5

<210> 335  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 335  
Ile Val Ser Gly Arg Ser Ser Xaa  
1 5

<210> 336  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 336  
Ile Val Ser Gly Arg Ala Ser Xaa  
1 5

<210> 337

-192-

<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 337  
Ile Val Ser Ala Arg Met Xaa  
1 5

<210> 338  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Xaa is Nle

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 338  
Ile Val Ser Ala Arg Xaa Xaa  
1 5

<210> 339  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 339  
Ile Val Ser Ala Arg Ser Xaa  
1 5

-193-

<210> 340  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nle-therapeutic Agent

<223> conjugate

<400> 340  
Ile Val Ser Ala Arg Ser Xaa  
1 5

<210> 341  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 341  
Ile Val Ser Ala Arg Ser Ser Xaa  
1 5

<210> 342  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> VARIANT  
<222> 8  
<223> dLeucine-therapeutic Agent

<223> conjugate

<400> 342  
Leu Arg Gly Ser Gly Arg Ser Xaa  
1 5

-194-

<210> 343  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 343  
Leu Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 344  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Nle-therapeutic agent

<223> conjugate

<400> 344  
Leu Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 345  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 345  
Leu Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 346

-195-

<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 346  
Leu Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 347  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 347  
Leu Arg Gly Ser Ala Arg Ser Xaa  
1 5

<210> 348  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 348  
Leu Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 349  
<211> 9

-196-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Nva-therapeutic agent

<223> conjugate

<400> 349  
Leu Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 350  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Nle-therapeutic agent

<223> conjugate

<400> 350  
Leu Arg Gly Ser Ala Arg Ser Ser Xaa  
1 5

<210> 351  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 351  
Val Ile Val Ser Gly Arg Ala Xaa  
1 5

<210> 352  
<211> 8  
<212> PRT

-197-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 352

Val Ile Val Ser Ala Arg Ser Xaa  
1 5

&lt;210&gt; 353

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 353

Val Ile Val Ser Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 354

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 354

Val Ile Val Ser Ala Arg Met Xaa  
1 5

&lt;210&gt; 355

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence



-198-

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Xaa is Nle

<221> MOD\_RES  
<222> 8  
<223> Alanine-therapeutic Agent

<223> conjugate

<400> 355  
Val Ile Val Ser Ala Arg Xaa Xaa  
1 5

<210> 356  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Nle-therapeutic agent

<223> conjugate

<400> 356  
Val Ile Val Ser Ala Arg Ser Xaa  
1 5

<210> 357  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> dCyclohexylalanine-therapeutic agent

<223> conjugate

<400> 357  
Val Ile Val Ser Ala Arg Ser Xaa  
1 5

<210> 358  
<211> 8

-199-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 358  
Val Ile Val Ser Ala Arg Ser Xaa  
1 5

<210> 359  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 9  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 359  
Val Ile Val Ser Ala Arg Ser Ser Xaa  
1 5

<210> 360  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is S-methylcysteine

<221> MOD\_RES  
<222> 9  
<223> Valine-therapeutic Agent

<223> conjugate

<400> 360  
Arg Arg Xaa Pro Gly Arg Val Xaa  
1 5

-200-

<210> 361  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is Nva

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 361  
Arg Arg Xaa Pro Ala Arg Ser Xaa  
1 5

<210> 362  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 362  
Ser Gly Arg Ser Xaa  
1 5

<210> 363  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 363  
Ser Gly Arg Ser Ser Xaa

-201-

1 5

<210> 364  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 364  
Ser Gly Arg Ser Ser Ser Xaa  
1 5

<210> 365  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 365  
Ser Gly Arg Ser Xaa  
1 5

<210> 366  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> dNva-therapeutic agent

<223> conjugate

<400> 366  
Ser Gly Arg Ser Xaa  
1 5

-202-

<210> 367  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nva-therapeutic agent  
  
<223> conjugate

<400> 367  
Ser Gly Arg Ser Xaa  
1 5

<210> 368  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 5  
<223> Hexylglycine-therapeutic agent  
  
<223> conjugate

<400> 368  
Ser Gly Arg Ser Xaa  
1 5

<210> 369  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent  
  
<223> conjugate

<400> 369  
Ser Gly Arg Ser Xaa  
1 5

-203-

<210> 370  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Homocyclohexylalanine-therapeutic agent  
  
<223> conjugate

<400> 370  
Ser Gly Arg Ser Xaa  
1 5

<210> 371  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent  
  
<223> conjugate

<400> 371  
Ser Ala Arg Ser Ser Xaa  
1 5

<210> 372  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent  
  
<223> conjugate

<400> 372  
Ser Ala Arg Ser Ser Xaa  
1 5

<210> 373

-204-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 373  
Ser Ser Arg Ser Xaa  
1 5

<210> 374  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Abu-therapeutic agent

<223> conjugate

<400> 374  
Thr Gly Arg Ser Xaa  
1 5

<210> 375  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 375  
Thr Gly Arg Ser Xaa  
1 5

<210> 376  
<211> 5

-205-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nva-therapeutic agent

<223> conjugate

<400> 376  
Thr Gly Arg Ser Xaa  
1 5

<210> 377  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 377  
Thr Gly Arg Ser Xaa  
1 5

<210> 378  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Hexylglycine-therapeutic agent

<223> conjugate

<400> 378  
Thr Gly Arg Ser Xaa  
1 5

<210> 379  
<211> 5  
<212> PRT



-206-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Cyclohexylalanine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 379

Thr Gly Arg Ser Xaa  
1 5

&lt;210&gt; 380

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Homocyclohexylalanine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 380

Thr Gly Arg Ser Xaa  
1 5

&lt;210&gt; 381

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Abu-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 381

Thr Gly Arg Thr Xaa  
1 5

&lt;210&gt; 382

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-207-

<220>  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 4  
<223> Xaa is Homoserine  
  
<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent  
  
<223> conjugate  
  
<400> 382  
Thr Gly Arg Xaa Xaa  
1 5

<210> 383  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 4  
<223> Xaa is Abu  
  
<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent  
  
<223> conjugate  
  
<400> 383  
Thr Gly Arg Xaa Xaa  
1 5

<210> 384  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 4  
<223> Xaa is Abu  
  
<221> MOD\_RES  
<222> 5  
<223> Nva-therapeutic agent  
  
<223> conjugate

-208-

<400> 384  
Thr Gly Arg Xaa Xaa  
1 5

<210> 385  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is 4-Guanidinophenylalanine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 385  
Thr Gly Xaa Ser Xaa  
1 5

<210> 386  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is 4-Guanidinophenylalanine

<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 386  
Thr Gly Xaa Ser Xaa  
1 5

<210> 387  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

-209-

<221> MOD\_RES  
<222> 3  
<223> Xaa is 4-Guanidinophenylalanine

<221> MOD\_RES  
<222> 4  
<223> Xaa is Abu

<221> MOD\_RES  
<222> 5  
<223> Nva-therapeutic agent

<223> conjugate

<400> 387  
Thr Gly Xaa Xaa Xaa  
1 5

<210> 388  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is Alloc

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 388  
Thr Gly Xaa Ser Xaa  
1 5

<210> 389  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 389  
Thr Gly Lys Ser Xaa  
1 5

-210-

<210> 390  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is Homoarginine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 390  
Thr Gly Xaa Ser Xaa  
1 5

<210> 391  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-homoserine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 391  
Xaa Gly Arg Ser Xaa  
1 5

<210> 392  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is N-Methyloxycarbonyl threonine

<221> MOD\_RES  
<222> 5

-211-

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 392

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 393

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Phenylsulfonyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 393

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 394

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Methoxyethylcarbonyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 394

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 395

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Methoxydiethoxyacetyl threonine

-212-

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent  
  
<223> conjugate

<400> 395  
Xaa Gly Arg Ser Xaa  
1 5

<210> 396  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is 4-Oxo-pentanoyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent  
  
<223> conjugate

<400> 396  
Xaa Gly Arg Ser Xaa  
1 5

<210> 397  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is 2-Benzo[1,3]dioxol-5-yl acetylthreonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent  
  
<223> conjugate

<400> 397  
Xaa Gly Arg Ser Xaa  
1 5

<210> 398  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1

-213-

&lt;223&gt; Xaa is 2-Pyridin-2-yl-acetyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 398

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 399

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Benzoylacetyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 399

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 400

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is 2-Hydroxy-3-phenyl propionylacetyl  
threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 400

Thr Gly Arg Ser Xaa  
1 5

&lt;210&gt; 401

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence



-214-

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxyacetylthreonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 401  
Xaa Gly Arg Ser Xaa  
1 5

<210> 402  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Phenylacetyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 402  
Xaa Gly Arg Ser Xaa  
1 5

<210> 403  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is 3-Methoxypropionyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 403  
Thr Gly Arg Ser Xaa  
1 5

<210> 404  
<211> 5  
<212> PRT

-215-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Methoxyethoxyacetyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 404

Thr Gly Arg Ser Xaa  
1 5

&lt;210&gt; 405

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is 1-Carboxybutanoyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 405

Thr Gly Arg Ser Xaa  
1 5

&lt;210&gt; 406

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Carboxybenzyl threonine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 406

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 407

-216-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Ethoxycarbonylthreonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 407  
Xaa Gly Arg Ser Xaa  
1 5

<210> 408  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is BALanine

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<223> conjugate

<400> 408  
Xaa Thr Gly Arg Ser Xaa  
1 5

<210> 409  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Pent-4-ynoyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 409  
Xaa Gly Arg Ser Xaa  
1 5

-217-

<210> 410  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Naphthaacetyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 410  
Xaa Gly Arg Ser Xaa  
1 5

<210> 411  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Isobutyloxycarbonyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 411  
Xaa Gly Arg Ser Xaa  
1 5

<210> 412  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Hydroxyacetyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 412  
Xaa Gly Arg Ser Xaa  
1 5

-218-

<210> 413  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarboxylpropanoyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 413  
Xaa Gly Arg Ser Xaa  
1 5

<210> 414  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is N,N-dimethyl glycine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 414  
Xaa Gly Arg Ser Xaa  
1 5

<210> 415  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Succinyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 415  
Xaa Gly Arg Ser Xaa

-219-

1 5

<210> 416  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> MOD\_RES  
<222> 1  
<223> Xaa is Formyl threonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 416  
Xaa Gly Arg Ser Xaa  
1 5

<210> 417  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 417  
Thr Ala Arg Ser Xaa  
1 5

<210> 418  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is 4-Guanidinophenylalanine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

-220-

&lt;223&gt; conjugate

&lt;400&gt; 418

Thr Ala Xaa Ser Xaa  
1 5

&lt;210&gt; 419

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Xaa is Abu

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nva-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 419

Thr Ala Arg Xaa Xaa  
1 5

&lt;210&gt; 420

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Abu-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 420

Thr Ala Arg Ser Xaa  
1 5

&lt;210&gt; 421

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

-221-

<222> 5  
<223> Abu-therapeutic agent

<223> conjugate

<400> 421  
Thr Ala Arg Thr Xaa  
1 5

<210> 422  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Serine methyl ester

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 422  
Thr Xaa Arg Ser Xaa  
1 5

<210> 423  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Homoserine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 423  
Thr Xaa Arg Ser Xaa  
1 5

<210> 424  
<211> 5  
<212> PRT



-222-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is 1-Methyl histidine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 424

Thr Xaa Arg Ser Xaa  
1 5

&lt;210&gt; 425

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is 3-Methyl histidine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 425

Thr Xaa Arg Ser Xaa  
1 5

&lt;210&gt; 426

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 426

-223-

Thr His Arg Ser Xaa  
1 5

<210> 427  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is MeGlycine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 427  
Thr Xaa Arg Ser Xaa  
1 5

<210> 428  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Nva

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 428  
Thr Xaa Arg Ser Xaa  
1 5

<210> 429  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

-224-

<221> MOD\_RES  
<222> 2  
<223> Xaa is Nle

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 429  
Thr Xaa Arg Ser Xaa  
1 5

<210> 430  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Cyclohexyl alanine-therapeutic agent

<223> conjugate

<400> 430  
Thr Ala Arg Ser Xaa  
1 5

<210> 431  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Abu

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 431  
Thr Xaa Arg Ser Xaa  
1 5

<210> 432  
<211> 5

-225-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is 4,4-Dimethylthreonine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 432  
Xaa Gly Arg Ser Xaa  
1 5

<210> 433  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is Homoserine

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 433  
Xaa Gly Arg Ser Xaa  
1 5

<210> 434  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1,4  
<223> Xaa is Homoserine

<221> MOD\_RES  
<222> 5

-226-

&lt;223&gt; Cyclohexyl alanine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 434

Xaa Gly Arg Xaa Xaa  
1 5

&lt;210&gt; 435

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Homoserine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Cyclohexylalanine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 435

Xaa Gly Arg Ser Xaa  
1 5

&lt;210&gt; 436

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Homoserine

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Cyclohexylalanine-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 436

Xaa Gly Arg Thr Xaa  
1 5

&lt;210&gt; 437

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-227-

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is Homoserine

<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 437  
Xaa Ala Arg Ser Xaa  
1 5

<210> 438  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Nle-therapeutic agent

<223> conjugate

<400> 438  
Asn Gly Arg Ser Xaa  
1 5

<210> 439  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 439  
Tyr Gly Arg Ser Ser Xaa  
1 5

<210> 440  
<211> 5

-228-

<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 440  
Tyr Gly Arg Ser Xaa  
1 5

<210> 441  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<223> conjugate

<400> 441  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 442  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<223> conjugate

<400> 442  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 443  
<211> 4  
<212> PRT

-229-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is N-homoserine

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Nle-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 443

Xaa Arg Ser Xaa

1

&lt;210&gt; 444

&lt;211&gt; 4

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is N-homoserine

&lt;221&gt; MOD\_RES

&lt;222&gt; 4

&lt;223&gt; Nva-therapeutic agent

&lt;223&gt; conjugate

&lt;400&gt; 444

Xaa Arg Ser Xaa

1

&lt;210&gt; 445

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; BLOCKED

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic Agent

&lt;223&gt; conjugate

&lt;400&gt; 445



-230-

Gln Gly Arg Ser Xaa  
1 5

<210> 446  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 446  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 447  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 447  
Gln Gly Arg Ala Ser Xaa  
1 5

<210> 448  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic Agent

<223> conjugate

<400> 448  
Asn Gly Arg Ser Ser Xaa

-231-

1

5

<210> 449  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Nleucine-therapeutic agent  
  
<223> conjugate

<400> 449  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 450  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nvaline-therapeutic agent

<223> conjugate

<400> 450  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 451  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<223> conjugate

<400> 451  
Gln Gly Arg Ser Ser Xaa  
1 5

-232-

<210> 452  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Allyl-therapeutic agent

<223> conjugate

<221> MOD\_RES  
<222> 5  
<223> Xaa is dSerine

<400> 452  
Gln Gly Arg Ser Xaa Xaa  
1 5

<210> 453  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Allyl-therapeutic agent

<400> 453  
Gln Gly Arg Ser Ser Xaa  
1 5

<210> 454  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 454

-233-

Gln Ala Arg Ser Xaa  
1 5

<210> 455  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 455  
Gln Ala Arg Ser Ser Xaa  
1 5

<210> 456  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 456  
Gln Ser Arg Ser Xaa  
1 5

<210> 457  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<400> 457  
Gln Ser Arg Ser Ser Xaa

-234-

1 5

<210> 458  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 458  
Gln Ser Arg Ser Ser Xaa  
1 5

<210> 459  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 459  
Gln Ser Arg Ser Ser Xaa  
1 5

<210> 460  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 460  
Gln Thr Arg Ser Ser Xaa  
1 5

-235-

<210> 461  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Aib

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 461  
Gln Xaa Arg Ser Ser Xaa  
1 5

<210> 462  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Aib

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 462  
Gln Xaa Arg Ser Ser Xaa  
1 5

<210> 463  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> (0) ... (0)

<221> MOD\_RES

-236-

<222> 2  
<223> Xaa is Abu  
  
<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 463  
Gln Xaa Arg Ser Ser Xaa  
1 5

<210> 464  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Abu

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 464  
Gln Xaa Arg Ser Ser Xaa  
1 5

<210> 465  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2  
<223> Xaa is Cyclohexylalanine

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 465  
Gln Xaa Arg Ser Ser Xaa  
1 5

<210> 466

-237-

<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 466  
Gln Phe Arg Ser Xaa  
1 5

<210> 467  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 467  
Gln Phe Arg Ser Ser Xaa  
1 5

<210> 468  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 468  
Gln Tyr Arg Ser Ser Xaa  
1 5

<210> 469  
<211> 5



-238-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 469  
Arg Gly Arg Ser Xaa  
1 5

<210> 470  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 470  
Arg Gly Arg Ser Ser Xaa  
1 5

<210> 471  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 471  
Arg Gly Arg Ser Ser Xaa  
1 5

<210> 472  
<211> 5  
<212> PRT

-239-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Cyclohexylalanine-therapeutic agent

&lt;400&gt; 472

Arg Gly Arg Ser Xaa  
1 5

&lt;210&gt; 473

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 473

Arg Ala Arg Ser Xaa  
1 5

&lt;210&gt; 474

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 474

Arg Ala Arg Ser Ser Xaa  
1 5

&lt;210&gt; 475

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-240-

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 475  
Arg Ser Arg Ser Xaa  
1 5

<210> 476  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 476  
Arg Ser Arg Ser Ser Xaa  
1 5

<210> 477  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent

<400> 477  
Arg Ser Arg Ser Xaa  
1 5

<210> 478  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>

-241-

<223> conjugate  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 478  
Arg Ser Arg Ser Ser Xaa  
1 5

<210> 479  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 479  
Arg Phe Arg Ser Xaa  
1 5

<210> 480  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Cyclohexylalanine-therapeutic agent

<400> 480  
Arg Phe Arg Ser Xaa  
1 5

<210> 481  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

-242-

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 481  
Tyr Gly Arg Ser Ser Xaa  
1 5

<210> 482  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is S-Dioxomethionine

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 482  
Xaa Ser Arg Ser Xaa  
1 5

<210> 483  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-cyanobenzyl) )  
glutamic acid-delta-methyl ester

<221> AMIDATION  
<222> 5

<400> 483  
Xaa Gly Arg Ser Leu  
1 5

<210> 484  
<211> 5  
<212> PRT  
<213> Artificial Sequence

-243-

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-amidinobenzyl) )  
glutamic acid -delta-methyl ester

<221> AMIDATION  
<222> 5

<400> 484  
Xaa Gly Arg Ser Leu  
1 5

<210> 485  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-  
(alpha-(3-amidinobenzyl)) glutamic acid

<221> AMIDATION  
<222> 5

<400> 485  
Xaa Gly Arg Ser Leu  
1 5

<210> 486  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-Methylbenzyl)  
)glutamic acid -delta-methyl ester

<221> AMIDATION  
<222> 5

<400> 486  
Xaa Gly Arg Ser Leu  
1 5

<210> 487  
<211> 5  
<212> PRT

-244-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is Methoxycarbonyl-  
(alpha-(3-methylbenzyl)) glutamic acid

&lt;221&gt; AMIDATION

&lt;222&gt; 5

&lt;400&gt; 487

Xaa Gly Arg Ser Leu  
1 5

&lt;210&gt; 488

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is Methoxycarbonyl-(alpha-(3-cyanobenzyl) )  
glutamic acid-delta-methyl ester

&lt;221&gt; AMIDATION

&lt;222&gt; 6

&lt;400&gt; 488

Xaa Gly Arg Ser Ser Leu  
1 5

&lt;210&gt; 489

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is Methoxycarbonyl-(alpha-(3-methylbenzyl)  
)glutamic acid -delta-methyl ester

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 489

Xaa Gly Arg Ser Xaa  
1 5

-245-

<210> 490  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-cyanobenzyl) )  
glutamic acid -delta-methyl ester

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 490  
Xaa Gly Arg Ser Xaa  
1 5

<210> 491  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLTATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 491  
Arg Gln Gly Arg Ser Xaa  
1 5

<210> 492  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLTATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 492  
Arg Gln Gly Arg Ser Ser Xaa  
1 5



-246-

<210> 493  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 493  
Arg Gln Gly Arg Ser Xaa  
1 5

<210> 494  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<400> 494  
Arg Gln Gly Arg Ser Xaa  
1 5

<210> 495  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Phenylalanine-therapeutic agent

<400> 495  
Arg Gln Gly Arg Ser Xaa  
1 5

<210> 496

-247-

<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> dLeucine-therapeutic agent  
  
<400> 496  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 497  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 497  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 498  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> dNle-therapeutic agent  
  
<400> 498  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 499  
<211> 6

-248-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 499  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 500  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<400> 500  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 501  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 501  
Arg Gln Gly Arg Ala Xaa  
1 5

<210> 502  
<211> 6  
<212> PRT

-249-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Phenylalanine-therapeutic agent

&lt;400&gt; 502

Arg Gln Gly Arg Ala Xaa  
1 5

&lt;210&gt; 503

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; leucine-therapeutic agent

&lt;400&gt; 503

Arg Asn Gly Arg Ser Xaa  
1 5

&lt;210&gt; 504

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nle-therapeutic agent

&lt;400&gt; 504

Arg Asn Gly Arg Ala Xaa  
1 5

&lt;210&gt; 505

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-250-

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 505  
Arg Gln Ala Arg Ser Xaa  
1 5

<210> 506  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 506  
Arg Gln Ala Arg Ser Xaa  
1 5

<210> 507  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<400> 507  
Arg Gln Ala Arg Ser Xaa  
1 5

<210> 508  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>

-251-

<223> conjugate  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 508  
Arg Gln Ala Arg Ser Xaa  
1 5

<210> 509  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7  
<223> Cyclohexylalanine-therapeutic agent

<400> 509  
Arg Gln Ala Arg Ser Ser Xaa  
1 5

<210> 510  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 510  
Arg Gln Ala Arg Thr Xaa  
1 5

<210> 511  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

-252-

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 511

Arg Gln Ala Arg Ala Xaa  
1 5

&lt;210&gt; 512

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nle-therapeutic agent

&lt;400&gt; 512

Arg Gln Ala Arg Ala Xaa  
1 5

&lt;210&gt; 513

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nva-therapeutic agent

&lt;400&gt; 513

Arg Gln Ala Arg Ala Xaa  
1 5

&lt;210&gt; 514

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

-253-

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Cyclohexylalanine-therapeutic agent

&lt;400&gt; 514

Arg Gln Ala Arg Ala Xaa

1

5

&lt;210&gt; 515

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 515

Arg Gln Ser Arg Ala Xaa

1

5

&lt;210&gt; 516

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 516

Arg Gln Ser Arg Xaa

1

5

&lt;210&gt; 617

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1



-254-

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 517  
Arg Gln Ser Arg Ala Xaa  
1 5

<210> 518  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 518  
Arg Gln Ser Arg Ala Xaa  
1 5

<210> 519  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 519  
Arg Gln Ser Arg Ala Xaa  
1 5

<210> 520  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES

-255-

<222> 6  
<223> Nva-therapeutic agent

<400> 520  
Arg Gln Ser Arg Ala Xaa  
1 5

<210> 521  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLTATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 521  
Arg Gln Ser Arg Ala Xaa  
1 5

<210> 522  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLTATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 522  
Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 523  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLTATION  
<222> 1

<221> MOD\_RES  
<222> 6

-256-

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 523

Arg Gln Ser Arg Ser Xaa  
1 5

&lt;210&gt; 524

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; dNle-therapeutic agent

&lt;400&gt; 524

Arg Gln Ser Arg Ser Xaa  
1 5

&lt;210&gt; 525

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nle-therapeutic agent

&lt;400&gt; 525

Arg Gln Ser Arg Ser Xaa  
1 5

&lt;210&gt; 526

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Nva-therapeutic agent

-257-

<400> 526  
Arg Gln Ser Arg Ser Xaa  
1 5

<210> 527  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Allylglycine-therapeutic agent

<400> 527  
Arg Gln Ser Arg Ser Xaa  
1 5

<210> 528  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 528  
Arg Gln Ser Arg Ser Xaa  
1 5

<210> 529  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle

<400> 529

-258-

Arg Gln Ser Arg Thr Xaa  
1 5

<210> 530  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 530  
Arg Gln Thr Arg Ser Ser Xaa  
1 5

<210> 531  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 531  
Arg Gln Thr Arg Ser Xaa  
1 5

<210> 532  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 532  
Arg Asn Ser Arg Ser Xaa

-259-

1

5

<210> 533  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 533  
Arg Gln Phe Arg Ser Xaa  
1 5

<210> 534  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> dNle-therapeutic agent  
  
<400> 534  
Arg Gln Phe Arg Ser Xaa  
1 5

<210> 535  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent  
  
<400> 535  
Arg Gln Phe Arg Ser Xaa  
1 5

-260-

<210> 536  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 536  
Arg Gln Phe Arg Ser Xaa  
1 5

<210> 537  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 537  
Arg Gln Phe Arg Ser Xaa  
1 5

<210> 538  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 538  
Arg Gln Phe Arg Ala Xaa  
1 5

-261-

<210> 539  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 539  
Arg Gln Phe Arg Ala Xaa  
1 5

<210> 540  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nva-therapeutic agent

<400> 540  
Arg Gln Phe Arg Ala Xaa  
1 5

<210> 541  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Cyclohexylalanine-therapeutic agent

<400> 541  
Arg Gln Phe Arg Ala Xaa  
1 5

<210> 542



-262-

<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Nle-therapeutic agent

<400> 542  
Gln Ser Arg Ser Ser Xaa  
1 5

<210> 543  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 543  
Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

<210> 544  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 544  
Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 545  
<211> 7

-263-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

/ <400> 545  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 546  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nle-therapeutic agent

<400> 546  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 547  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Nle-therapeutic agent

<400> 547  
Arg Gly Ser Gly Arg Ala Xaa  
1 5

<210> 548  
<211> 8  
<212> PRT

-264-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 548

Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 549

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; (0)...(0)

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 549

Ile Val Ser Gly Arg Ala Ser Xaa  
1 5

&lt;210&gt; 550

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 550

Leu Arg Arg Gln Ser Arg Ser Ser Xaa  
1 5

&lt;210&gt; 551

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-265-

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent  
  
<400> 551  
Leu Arg Arg Gln Ser Arg Ser Xaa  
1 5

<210> 552  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 552  
Gln Ser Arg Ala Xaa  
1 5

<210> 553  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Alanine-therapeutic agent

<400> 553  
Gln Ser Arg Ser Xaa  
1 5

<210> 554  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>

-266-

&lt;223&gt; conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Glycine-therapeutic agent

&lt;400&gt; 554

Gln Ser Arg Ser Xaa  
1 5

&lt;210&gt; 555

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 555

Arg Ser Arg Ala Xaa  
1 5

&lt;210&gt; 556

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 556

Arg Gln Ser Arg Ala Xaa  
1 5

&lt;210&gt; 557

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

-267-

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alalnine-therapeutic agent

&lt;400&gt; 557

Arg Gln Ser Arg Ser Xaa  
1 5

&lt;210&gt; 558

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 558

Arg Gln Ser Arg Ser Ala Xaa  
1 5

&lt;210&gt; 559

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 559

Arg Gly Ser Gly Arg Ser Xaa  
1 5

&lt;210&gt; 560

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

-268-

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 560

Ser Gly Arg Ala Xaa  
1 5

&lt;210&gt; 561

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 561

Ser Gly Arg Ser Xaa  
1 5

&lt;210&gt; 562

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 562

Ser Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 563

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

-269-

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 563  
Ser Gly Arg Ala Ser Xaa  
1 5

<210> 564  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Glycine-therapeutic agent

<400> 564  
Ser Gly Arg Ser Xaa  
1 5

<210> 565  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Glycine-therapeutic agent

<400> 565  
Ser Gly Arg Ser Ser Xaa  
1 5

<210> 566  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES



-270-

<222> 6  
<223> Alanine-therapeutic agent

<400> 566  
Ser Gly Arg Ser Gly Xaa  
1 5

<210> 567  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Glycine-therapeutic agent

<400> 567  
Ser Gly Arg Ser Gly Xaa  
1 5

<210> 568  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Glycine-therapeutic agent

<400> 568  
Gly Thr Gly Arg Ser Gly Xaa  
1 5

<210> 569  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 2

-271-

&lt;223&gt; Xaa is D- Serine

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 569

Gly Xaa Ala Arg Ser Xaa  
1 5

&lt;210&gt; 570

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 3

&lt;223&gt; Xaa is D-Serine

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 570

Arg Gly Xaa Ala Arg Ser Xaa  
1 5

&lt;210&gt; 571

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 571

Gly Ser Gly Arg Ser Xaa  
1 5

&lt;210&gt; 572

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

-272-

<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent

<400> 572  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 573  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Alanine-therapeutic agent

<400> 573  
Leu Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 574  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 4  
<223> Xaa is D-Serine

<221> MOD\_RES  
<222> 8  
<223> Alanine-therapeutic agent

<400> 574  
Leu Arg Gly Xaa Ala Arg Ser Xaa  
1 5

<210> 575  
<211> 6  
<212> PRT

-273-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is S-Methylcysteine

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Valine-therapeutic agent

&lt;400&gt; 575

Xaa Pro Gly Arg Val Xaa

1

5

&lt;210&gt; 576

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is S-Methylcysteine

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Valine-therapeutic agent

&lt;400&gt; 576

Xaa Pro Gly Arg Val Xaa

1

5

&lt;210&gt; 577

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is S-Methylcysteine

&lt;221&gt; MOD\_RES

-274-

<222> 7  
<223> Valine-therapeutic agent

<400> 577  
Arg Xaa Pro Gly Arg Val Xaa  
1 5

<210> 578  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 3  
<223> Xaa is S-Methylcysteine

<221> MOD\_RES  
<222> 8  
<223> Valine-therapeutic agent

<400> 578  
Arg Arg Xaa Pro Gly Arg Val Xaa  
1 5

<210> 579  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent

<400> 579  
Val Ser Ala Arg Met Xaa  
1 5

<210> 580  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION

-275-

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 580

Ile Val Ser Ala Arg Met Xaa  
1 5

&lt;210&gt; 581

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 581

Val Ile Val Ser Ala Arg Met Xaa  
1 5

&lt;210&gt; 582

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 7

&lt;223&gt; Xaa is Nle

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 582

Val Ile Val Ser Ala Arg Xaa Xaa  
1 5

&lt;210&gt; 583

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

-276-

<223> conjugate  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Xaa is Nle  
  
<221> MOD\_RES  
<222> 6  
<223> Alanine-therapeutic agent  
  
<400> 583  
Val Ser Ala Arg Xaa Xaa  
1 5

<210> 584  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Xaa is Nle  
  
<221> MOD\_RES  
<222> 7  
<223> Alanine-therapeutic agent  
  
<400> 584  
Ile Val Ser Ala Arg Xaa Xaa  
1 5

<210> 585  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 585  
Gly Ser Gly Arg Ser Xaa  
1 5

-277-

<210> 586  
<211> 7  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent  
  
<400> 586  
Gly Ser Gly Arg Ser Ser Xaa  
1 5  
  
<210> 587  
<211> 6  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent  
  
<400> 587  
Gly Ser Ala Arg Ser Xaa  
1 5  
  
<210> 588  
<211> 5  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> conjugate  
  
<221> ACETYLATION  
<222> 1  
  
<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent  
  
<400> 588  
Ser Gly Arg Ser Xaa  
1 5  
  
<210> 589



-278-

<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 589  
Ser Gly Arg Ser Ser Xaa  
1 5

<210> 590  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 590  
Ser Ala Arg Ser Xaa  
1 5

<210> 591  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 591  
Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 592  
<211> 8

-279-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 592  
Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

<210> 593  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 7  
<223> Leucine-therapeutic agent

<400> 593  
Arg Gly Ser Ala Arg Ser Xaa  
1 5

<210> 594  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 8  
<223> Leucine-therapeutic agent

<400> 594  
Leu Arg Gly Ser Gly Arg Ser Xaa  
1 5

<210> 595  
<211> 9  
<212> PRT

-280-

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 9

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 595

Leu Arg Gly Ser Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 596

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 596

Leu Arg Gly Ser Ala Arg Ser Xaa  
1 5

&lt;210&gt; 597

&lt;211&gt; 8

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Alanine-therapeutic agent

&lt;400&gt; 597

Leu Arg Arg Gln Ser Arg Ala Xaa  
1 5

&lt;210&gt; 598

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-281-

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is N-Methylsulfonyl-alpha-cyclohexyl-D-Alanine

<221> MOD\_RES  
<222> 2  
<223> Xaa is Abu

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 598  
Xaa Xaa Arg Ser Xaa  
1 5

<210> 599  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 5  
<223> Leucine-therapeutic agent

<400> 599  
Arg Ala Arg Ser Xaa  
1 5

<210> 600  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa si Alpha-Cyclohexyl-D-alanine

<221> MOD\_RES  
<222> 2  
<223> Abu

<221> MOD\_RES  
<222> 5

-282-

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 600

Xaa Xaa Arg Ser Xaa  
1 5

&lt;210&gt; 601

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Alpha-Cyclohexyl-D-Alanine

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Abu

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 601

Xaa Xaa Arg Ser Ser Xaa  
1 5

&lt;210&gt; 602

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 602

Gln Gly Arg Ser Ser Xaa  
1 5

&lt;210&gt; 603

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

-283-

&lt;223&gt; conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is Methoxycarbonyl-D-homophenylalanine

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is 4Hyp

&lt;221&gt; MOD\_RES

&lt;222&gt; 6

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 603

Xaa Xaa Arg Ser Ser Xaa

1

5

&lt;210&gt; 604

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

<223> Xaa is Methoxycarbonyl-(alpha)-3-methylbenzyl  
glutamic acid -delta-methyl ester

&lt;221&gt; MOD\_RES

&lt;222&gt; 5

&lt;223&gt; Leucine-therapeutic agent

&lt;400&gt; 604

Xaa Gly Arg Ser Xaa

1

5

&lt;210&gt; 605

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; conjugate

&lt;221&gt; ACETYLATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 1

&lt;223&gt; Xaa is D-cyclohexylalanine

&lt;221&gt; MOD\_RES

&lt;222&gt; 2

&lt;223&gt; Xaa is 4Hyp

&lt;221&gt; MOD\_RES

-284-

<222> 5  
<223> Leucine-therapeutic agent

<400> 605  
Xaa Xaa Arg Ser Ser Xaa  
1 5

<210> 606  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> ACETYLATION  
<222> 1

<221> MOD\_RES  
<222> 1  
<223> Xaa is D-Clohexylalanine

<221> MOD\_RES  
<222> 2  
<223> Xaa is Abu

<221> MOD\_RES  
<222> 6  
<223> Leucine-therapeutic agent

<400> 606  
Xaa Xaa Arg Ser Ser Xaa  
1 5

<210> 607  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-cyanobenzyl))  
glutamic acid -delta-methyl ester

<221> AMIDATION  
<222> 5

<400> 607  
Xaa Gly Arg Ser Leu  
1 5

<210> 608  
<211> 5  
<212> PRT  
<213> Artificial Sequence

-285-

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-amidinobenzyl))  
glutamic acid -delta-methyl ester

<221> AMIDATION  
<222> 5

<400> 608  
Xaa Gly Arg Ser Leu  
1 5

<210> 609  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> conjugate

<221> MOD\_RES  
<222> 1  
<223> Xaa is Methoxycarbonyl-(alpha-(3-amidinobenzyl))  
glutamic acid

<221> AMIDATION  
<222> 5

<400> 609  
Xaa Gly Arg Ser Leu  
1 5

<210> 610  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Conjugate

<221> ACETYLATION  
<222> (0)...(0)

<221> MOD\_RES  
<222> 6  
<223> Isoleucine-therapeutic agent

<400> 610  
Arg Arg Gln Ser Arg Xaa  
1 5

<210> 611  
<211> 8  
<212> PRT  
<213> Artificial Sequence



-286-

&lt;220&gt;

&lt;223&gt; Conjugate

&lt;221&gt; ACETYLTATION

&lt;222&gt; 1

&lt;221&gt; MOD\_RES

&lt;222&gt; 8

&lt;223&gt; Isoleucine-therapeutic agent

&lt;400&gt; 611

Leu Arg Arg Gln Ser Arg Ala Xaa  
1 5